What Does Governance Do? Evidence From Airlines*

Silke Forbes† Ricard Gil‡ Giorgio Zanarone§

December 30, 2025

Abstract

Extant research on firm boundaries provides limited insight on the mechanisms through which vertical integration improves supplier performance, and on the extent to which contracts can substitute for it. Using rich data from the US airline industry, we shed new light on both questions. First, we show that flights ticketed by a major airline are less delayed if operated by the major's employees than by the employees of an outsourced regional partner, holding the major's control over alienable decisions constant. This evidence provides rare support for the view of the firm as a mechanism to monitor and incentivize employees (Holmstrom and Milgrom, 1991; Holmstrom, 1999). Second, we compare vertical integration to relational and arm's length outsourcing, and find that both vertically integrated flights and flights outsourced to relational partners experience similar delay reductions compared to flights outsourced to arm's length partners. This results suggests that comparing vertical integration to an indistinct bundle of outsourcing relationships may overstate its effect.

JEL Codes: D22, D23, L14, L24, L93

*We acknowledge the financial support obtained from the Swiss National Science Foundation through grant 10.001.670. We are grateful to Riccardo Camboni, Wouter Dessein, Florian Englmaier, Rachel Kranton, Brian Silverman, Pablo Spiller, and audiences at HEC Lausanne, IESE, and the Paris-Berkely Organizational Economics Workshop, for useful comments.

†Tufts University, silke.forbes@tufts.edu

‡IESE Business School, rgils@iese.edu

§HEC Lausanne, University of Lausanne, giorgio.zanarone@unil.ch

1 INTRODUCTION

Ever since the foundational work of Coase (1937) and Williamson (1971), a central question in organizational economics and strategy has been the make-or-buy decision: do firms obtain better inputs when they produce them internally or when they buy them from outsourced providers? Evidence from several industries has convincingly shown that vertical integration emerges when competitive spot markets for inputs are unavailable, and thus buyers cannot credibly threaten to punish poorly performing suppliers by switching to a competitor (see Lafontaine and Slade (2007) for a review). Several decades later, however, we still have limited understanding of how vertical integration improves input provision and when it is truly necessary, in the sense that its outcome cannot be replicated by sophisticated buyer-supplier contracts.

The "how" question arises because there are several theoretical models of firm boundaries, each highlighting a different mechanism. These models can be grouped into two streams.[1] According to a "decision control view," vertical integration transfers the right to make alienable production decisions (what to supply and to whom) to the buyer. Decision rights may then enable the buyer to adapt inputs to unforeseen contingencies (Williamson, 1971, 1991) or to bargain with the upstream unit over such adaptations from a position of strength (Grossman and Hart, 1986; Hart and Moore, 1990). According to an "incentives view" of firm boundaries, the key decisions that determine input quality are inalienably delegated to agents (how much effort to exert and on what tasks) and thus cannot be transferred through vertical integration. However, integration enables the buyer to monitor, incentivize and regulate the supplier unit's employees in a way that would not be feasible under outsourcing (Masten, 1988; Holmstrom and Milgrom, 1991; Holmstrom, 1999).

[1]See Gibbons (2005a) and Dessein (2014) for integrative discussions of models of firm boundaries.

A few studies have exploited convenient industry features to show that vertical integration transfers alienable decisions to the buyer, such as a major airline delaying and canceling flights (Forbes and Lederman, 2009, 2010) or a fish processing firm deciding which plants fishing boats can sell to (Hansman et al., 2020). In contrast, there is very little evidence on integration as a tool to manage and incentivize employees, and the little evidence that exists is consistent with a negative effect of vertical integration on incentives (Baker and Hubbard, 2004). This paucity of evidence is not surprising because both the decision control and the incentive mechanisms are simultaneously activated by the transfer of upstream assets to the buyer and thus are hard to disentangle. Ownership of the supplier unit's assets enables the buyer both to take control of all alienable decisions over the use of those assets (Grossman and Hart, 1986) and to assign upstream employees to roles, monitor them, and strengthen or mute their incentives as desired (Holmstrom, 1999).

Regarding the "when" question (can contracts substitute for integration?), there is a vast literature, both theoretical and empirical, on how sophisticated formal contracts and self-enforcing relational contracts govern buyer-supplier relationships (see, among others, the reviews in Bolton and Dewatripont (2004); Lafontaine and Slade (2013); Gil and Zanarone (2017, 2018). However, empirical studies of firm boundaries have largely ignored the plurality of contractual forms available to govern suppliers, comparing instead vertically integration to an indistinct bundle of arm's length and relational outsourcing (Gibbons, 2025). Such bundling makes it difficult to interpret any observed positive effects of vertical integration on supplier behavior or firm performance because one cannot tell whether these effects are driven by the superior effectiveness of integration with respect to all contracts or by the prevalence of arm's length, unsophisticated contracts in the outsourcing subsample.

Our paper uses novel data from the US airline industry to shed light on both the understudied questions discussed above: (1) disentangling the incentive view from the decision control view of firm boundaries and (2) comparing the effectiveness of integration and buyer-supplier contracts. Both decision control and employee motivation are important and well-defined problems in the delivery of timely air transportation to passengers, and both problems are affected by the extent to which flights are operated or managed by the major (vertical integration) or by an independent regional partner (outsourcing) and by the nature of outsourcing relationships. On the decision side, operating a flight gives the major the right to use its landing slot to flexibly reschedule other flights – a frequent necessity during adverse weather (Forbes and Lederman, 2009, 2010; Gil et al., 2022). On the incentives side, employing a flight's pilots, ground managers and crew enables the major to monitor and incentivize those agents to exert the effort necessary to ensure on-time performance.

To separate the incentive effect of vertical integration from the decision-control effect, we exploit the fact that while a major's ability to manage employees in charge of a flight depends on who operates that flight (flight integration), the major's ability to reduce a flight's delay through slot exchanges and rescheduling only depends on who operates other flights serving the same route (Forbes and Lederman, 2010). Thus, by comparing the on-time performance of similar vertically integrated and outsourced flights of a given major on the same route, we can estimate the performance effect of vertical integration coming from the improved management of flight employees (incentive view), while holding the impact of slot control (decision view) constant. We find that which firm manages the employees in charge of a flight strongly affects that flight's performance: the average delay of flights operated by the major is about seven minutes lower than the delay of otherwise similar outsourced flights, and the difference becomes even more pronounced in adverse weather days in which the employees' effort matters more at the margin.

To compare the effect of vertical integration and different forms of outsourcing contracts on performance, we construct measures of the present discounted value of major-regional relationships using the methodology developed by Gil et al. (2022). Consistent with standard relational contracting theory (Baker et al., 2002; Levin, 2003), we find that when the regional partner operating a flight has a high-value relationship with the major (and hence the major can fully rely on self-enforcing agreements to ensure efficient people management by the regional), the delay of outsourced flights is similar to that of integrated flights. The performance advantage of vertically integrated flights over outsourced ones therefore appears to be driven by low-value outsourcing relationships. These findings suggest that major airlines may use their long-term informal relationships with certain regional partners to ensure those partners' employees are managed as if they were the major's.

The rest of this paper is organized as follows. Section 2 discusses our contributions to the literature. Section 3 briefly describes buyer-supplier relationships in the US airline industry. Section 4 presents a simple model of people management under vertical integration and outsourcing, which generates testable predictions. Section 5 presents our data and empirical methodology. Section 6 discusses the results. Section 7 concludes.

2 RELATION TO THE LITERATURE

2.1 Firm Boundaries

Our paper contributes to both the literature on firm boundaries and the literature on relational contracting. As discussed above, there is abundant evidence that consistent with classic transaction cost economics (Coase, 1937; Williamson, 1971, 1979, 1991), vertical integration replaces outsourcing when spot markets for inputs are unavailable (e.g., Monteverde and Teece (1982); Masten (1984). This classic literature has been extensively reviewed by Lafontaine and Slade (2007). Evidence on the mechanisms through which vertical integration helps is more limited and mostly related to the "decision control" view of firm boundaries. There is evidence that as argued by Grossman and Hart (1986), vertical integration transfers decision rights to the buyer, improving her bargaining position and thus encouraging her specific investments at the expense of the supplier's (Acemoglu et al., 2010). There is also evidence that by reducing the supplier's discretion to make decisions, vertical integration may reduce her ability to pursue private benefits at the expense of service quality (Hansman et al., 2020).

Closer to our study, a few papers on the US airline industry (Forbes and Lederman, 2009, 2010) have shown that integration enables the buyer to make non-contractible decisions that an independent supplier would be unwilling to make, as argued by Williamson (1971, 1991). Forbes and Lederman (2010) is perhaps the closest to our paper as they investigate the effect of vertical integration on flights' on-time performance, finding that major-operated flights experience shorter delays at airports where the major is more vertically integrated. In contrast, our paper holds the major's control of airport decisions constant and compares the performance of vertically integrated and outsourced flights, shedding light on the incentive view of firm boundaries that the existing empirical literature has largely ignored.

To the best of our knowledge, the only paper that examines the effect of vertical integration on incentive provision is Baker and Hubbard (2004). Using data on US trucking, they show that when on-board computers enable shipping companies to monitor truck drivers, vertical integration (trucks owned by the shipper) increases relative to outsourcing (trucks owned by the driver). This evidence is consistent with the view that when an agent does not own the assets she operates, and hence is not residual claimant of their value, she has weaker incentives to preserve them – that is, vertical integration mutes incentives (Holmstrom and Milgrom, 1991). Our study differs from Baker and Hubbard (2004) in two fundamental ways. First, it provides evidence that vertical integration enhances agents' performance, rather than decreasing it. Second, in our setting the employees affecting flight performance do not own any productive assets under either vertical integration or outsourcing. Thus, the effect on vertical integration we observe is inconsistent with a change in employees' residual claimant rights. Instead, our evidence is consistent with the idea that who is the employer, which firm sets the employment rules and incentives, critically affects the performance of employees (Holmstrom, 1999).

2.2 Relational Contracts

An empirical literature flourished in the last decade has analyzed the determinants and consequences of relational contract – self-enforcing agreements sustained by repeated interactions between the parties (Baker et al., 2002). Papers in this literature have shown that the behavior of contracting parties critically depends on industry, transaction and relationship features that affect the long-run gains from cooperation and short-term gains from defection. For instance, Macchiavello and Morjaria (2015) have shown that shocks to the reneging temptation of flower exporters (increases in the spot market price) leads to a reduction in their shipments to relational clients that keeps the relationship self-enforcing. Barron et al. (2020) have shown that movie distributors grant voluntary and non-contractual discounts to exhibitors who keep showing their movies even after more profitable alternatives are available. Macchiavello and Morjaria (2021) have shown that in the Rwanda coffee industry, competition between mills reduces the value of their relationships with farmers and hence the use of relational contracts to govern farmer-mill relationships. Gil et al.

(2022) have shown that in the US airline industry, outsourcing relationships with higher continuation value are more resilient to a systemic shock (the 2008 financial crisis) that calls for the suppression of routes. Gil and Zanarone (2017, 2018), and Macchiavello (2022), offer comprehensive reviews of this literature.

While this paper builds on the methodology used by Gil et al. (2022) to measure the value of major-regional outsourcing relationships, it innovates on that paper – and on all the other papers discussed above – in that instead of restricting attention to interfirm relationships, it compares the effect of relational contracts between firms and vertical integration on performance, relative to arm's length contractual relationships. The only other paper we are aware of that conducts this kind of horserace between governance forms is Macchiavello and Miquel-Florensa (2019), which compares vertically integrated relationships in the Costa Rican coffee chain to outsourcing relationships characterized by many versus few past interactions. Their paper differs from ours in that it studies a context in which vertical integration transfers decision control rather than employee incentives, and explores relational contracts as an alternative, informal tool to do so.

3 VERTICAL RELATIONSHIPS IN THE US AIRLINE INDUSTRY

3.1 Contracts and Firm Boundaries

In the U.S. airline industry, major airlines (e.g., United Airlines, Delta, American Airlines, Alaska Airlines) operate larger aircraft to serve (mostly) long-distance routes. Regional airlines operate smaller aircraft to serve shorter routes connecting a hub to a local airport, and are sometimes owned by the major but most often they are independent companies that partner with the major to operate flights under the latter's banner. Examples of major-owned regionals are Piedmont Airlines or Endeavor Air. Examples of independent partner regionals are Skywest, Air Wisconsin, Trans States Airlines, and Republic Airways.

Relationships between majors and regional partners (both integrated and outsourced) are governed by "capacity purchase agreements" (CPAs), under which the regional operates specific flights under the major's brand, while the major sets flight schedules, sells tickets, and buys fuel for such flights. The major collects all revenues, and the regional receives a flat fee for each operated flight. CPAs also include provisions providing the regional partner with some incentives for timely performance, such as minimum quotas of timely landings that are necessary to collect the contractual fees, and bonuses linked to on-time performance.

The primary purpose of outsourcing agreements is to reduce the cost of operating connection flights, especially because independent regional airlines can avoid paying the higher, union-negotiated wages and benefits to pilots, flight attendants, and mechanics that majors must pay (Forbes and Lederman, 2009). It is important to stress that CPAs are by no means exclusive agreements – a given major typically relies on multiple regional partners, even on the same route, and regional airlines are often large companies that serve multiple majors. Moreover, while formal CPAs are boilerplate and very similar across major-regional pairs, there is considerable variation, even within a major, in the length and value of outsourcing relationships, and thus in the major's ability to enter relational, self-enforcing agreements with its partner (Gil et al., 2022).

Both alienable and inalienable decisions are important for service quality in the US airline industry. These decisions are affected by (a) whether a flight is operated by the major or by an outsourced regional and (b) in the event of outsourcing, by the strength of the underlying relationship.

3.2 Alienable Decisions

The key alienable decision in this industry is the use of landing slots to adapt flight schedules. Schedules at an airport are designed by major airlines months in advance, constituting an almost perfect plan to ensure the timely transportation of passengers between two airports anywhere in the USA. Unfortunately, these plans are often disrupted by inclement weather conditions, which may prompt the Federal Aviation Administration (FAA) to reduce the number of authorized landings through a Ground Delay Program (GDP). GDPs are ordered when safety demands such alteration of normal airport operations and they are very frequent events, occurring in almost three quarter of days during winter months at the more congested airports. To minimize landing delays on the most valuable flights affected by a GDP, major airlines seek to delay or cancel other flights scheduled to land at the airport to make a slot available for the focal flight (Argyres et al., 2025).

While the major owns the slots of flights it operates (directly or through a regional subsidiary) and can swap them around at will, it must obtain the regional partner's agreement to delay or cancel an outsourced flight. Such agreements must be reached ex post under severe time pressure as the difficulty of predicting GDPs prevents the inclusion of state-contingent slot exchange provisions in CPAs. Moreover, the regional has limited incentives to give up a slot because by doing so, it incurs personnel, logistics, and coordination costs while the benefits accrue to the ticketing airline – that is, the major (Forbes and Lederman, 2009; Argyres et al., 2025). Vertical integration therefore facilitates slot exchanges, relative to outsourcing, by transferring control over slots from the regional to the major.

3.3 Inalienable Decisions

Apart from the use of slots, the key decisions that affect a flight's performance are delegated to airline employees, and hence their control is not transferred to the major's upper management when the flight is vertically integrated. Some of these inalienable decisions – allocating scarce gates to early or late flights, directing ground crew teams that guide the flight to the gate and provide the tug at departure – are always made by the major's employees. Other decisions – allocating flight attendants and crew, directing planes out of the gate at departure – are made by employees of the operating airline, who are therefore monitored and managed by the major under vertical integration, and by a regional partner under outsourcing.

All these decisions affect a flight's arrival delay, especially at busy and congested airports and in days where there are multiple schedule disruptions and thus airport resources are scarce, such as when a GDP is enacted. For instance, conversations with industry practitioners indicate that regional planes occasionally fail to vacate departure gates as quickly as the major would have liked, presumably due to low effort of the regional employees in charge of staffing and directing the plane, which may result in both late departure and late arrival. Whether key employees at the operating airlines are properly monitored and incentivized by their employer is therefore important to the flight's on-time performance, and it potentially depends on whether the employer is the major or a regional partner, as shown below.

4 MODEL

Consider a flight ticketed by major airline M. The flight's value to M, $y = fa$, is determined by the effort $a \in \mathfrak{R}^{+}$ exerted by a representative agent A (a ground manager, a pilot) at cost $c = \frac{a^2}{2}$. This agent may be employed by M (integration) or by an independent regional partner R (non-integration). All players are risk-neutral and we assume for now that they interact only once. We consider repeated interactions in section 4.3 below.

The agent's efficient (first best) effort maximizes $E[p] - c = fa - \frac{a^2}{2}$, and is given by $a^{FB} = f$. The question we ask is which contractual arrangement gets M closest to reaching this efficient benchmark. We consider both input-based contracts, which condition payments directly on A's effort and thus require monitoring, and output-based contracts, which condition payments on some realized outcome that depends on effort.

Regarding output contracts, we following standard agency models in assuming that A's individual contribution to M's profit, y, is unverifiable by courts and hence non-contractible. However, a contractible but noisy proxy for A's effort exists and is given by the flight's on-time performance $p = fa + \epsilon$, where ϵ is a zero-mean random variable. We also follow the literature in assuming that incentive contracts based on p are generally inefficient. While there are many potential inefficiencies in output contracts, we focus for simplicity on limited liability by assuming that A and R cannot make upfront payments. As formally shown below, limited liability implies that output-based incentive contracts must leave a rent to the agent, leading the principal to elicit effort below a^{FB} (Sappington, 1983). It will soon become clear that modeling alternative frictions in output-based contracts, such as gaming (Holmstrom and Milgrom, 1991), would generate qualitatively similar results.

Input contracts clearly do better than output-based ones: an enforceable promise to pay A a salary equal to the effort cost $\frac{a^2}{2}$ if and only if she exerts effort a would elicit such effort without leaving a rent to the agent. In practice, monitoring and rules defining job assignments may reduce A's discretion to choose effort and implement input contracts. However, we follow Holmstrom (1999) in assuming that the ability to monitor and set rules crucially depends on who employs A: while M can enter an input contract with A under integration, only R can enter such contract under non-integration because an employer has exclusive monitoring and contracting rights over its employees (see also Masten (1988) and Williamson (1991)).[2] Thus, all M can do under non-integration is offer an output contract to R in the hope to influence R's input contract with A, leading to inefficiencies that we analyze below.

4.1 Managing Employees Under Integration

Under vertical integration, M can monitor A and contract with her on effort. We assume M can make a take-it-or-leave-it offer to A, implying that for a given effort a, it is optimal for M to offer a salary equal to the effort cost $c = \frac{a^2}{2}$. M will therefore select the effort level that solves

$$max_a \quad \{y - c = fa - \frac{a^2}{2}\},$$

which results in the first best effort $a^{VI} = f = a^{FB}$.

Notice that our model deliberately abstracts from potential costs of integration emphasized in the literature, such as distortions in non-contractible specific investments (Grossman and Hart, 1986; Hart and Moore, 1990). As discussed elsewhere (Forbes and Lederman, 2009; Argyres et al., 2025), specific investments are not a key concern in the airline industry.

4.2 Managing Employees Under Non-Integration

Under non-integration, R can monitor A and contract with her on effort, whereas M can only contract with R on the verifiable output p. We focus on linear incentive

[2] The employer's exclusive rights also imply that M cannot offer an output contract to A under non-integration. It will soon become clear, however, that even if feasible, such a contract would not add value.

contracts of the form $w = bp$, where w is the payment that M offers to R and b is a commission rate.[3] Given w, R designs the input contract with A to solve

$$max_a \quad \{E[w] - c = bfa - \frac{a^2}{2}\},$$

which results in the effort schedule $a^*(b) = bf$. Anticipating R's choice, M designs the output contract to solve

$$max_b \quad \{y - E[w] = fa^*(b) - bfa^*(b) = bf^2 - b^2 f^2\},$$

which results in the output sharing rule $b^* = 1/2$ that elicits effort $a^{NI} = a^*(b^*) = \frac{f}{2} < a^{FB}$. The non-integration effort is lower than the integration one because non-integration constrains the set of contractual instruments M can use to elicit effort. Being unable to monitor and contract with A, M cannot use input contracts. Instead, M must use an output contract to incentivize R to manage A. Given R's limited liability, in order to elicit a given effort a, this output contract must leave R a rent above the cost of effort, which in turn induces M to elicit less effort than would be efficient.[4].

The key takeaway is that in this model, integration increases effort and the flight's value, relative to non-integration, not because it transfers control over the effort decision to M (it is always A who chooses effort, regardless of firm boundaries), as in decision-control theories of the firm (Grossman and Hart, 1986; Hart and Moore, 1990; Williamson, 1971). Instead, integration increases effort because it allows M

[3]Non-linear contracts may overcome the limited liability problem under certain conditions (Innes, 1990). However, non-linear contracts may introduce other distortions that render them impractical (Gibbons, 2005b). Moreover, linear contracts have been shown to be optimal in certain settings (Holmstrom and Milgrom, 1987).

[4]From the effort schedule $a^*(b) = bf$ we can compute the piece rate necessary to elicit a given effort a, which is $b(a) = \frac{a}{f}$. The compensation M must pay A to elicit effort a is therefore $b(a)E[p] = a^2$, which is twice the effort cost

to use incentive tools (monitoring and input contracts) that would not be available under non-integration (Masten, 1988; Holmstrom and Milgrom, 1991; Williamson, 1991; Holmstrom, 1999).

4.3 Relational Contracts

In this section we extend the baseline model to explore how relational contracts sustained by repeated interactions between M and R can make the non-integration equilibrium more efficient. To do so, we assume M and R live forever and discount next-period payoffs by the common factor $\delta \in (0, 1]$. We study subgame perfect equilibria of the repeated game that are stationary, in the sense that effort and payments are the same in every period. As discussed below, restricting attention to stationary contracts is without loss of optimality in our setting (Levin, 2003).

Relational contracts may help because while the output y is not verifiable, it is observed by M and R and thus relational contracts conditional on y are potentially feasible, provided the parties are patient enough. Since y is a perfect signal of A's effort, relational contracts may enable M to condition incentives on effort (rather than on the noisy performance measure p) and thus to avoid paying R a rent.

Formally, suppose that instead of offering a formal incentive contract based on p (as in section 4.2 above), M promises in each period to pay R a discretionary bonus b if and only if R instructs A to exert an informally agreed effort level a. Also, let $s > 0$ be a fixed salary that M pays R upfront in each period.[5] If R deviates on the promised effort or M deviates on the promised bonus, both parties revert to static non-integration forever after.[6]

[5]In theory, M could combine the relational bonus with a formal piece rate contract based on p, as in Baker et al. (1994). This richer contract would increase efficiency but would not generate additional insights so we ignore it for simplicity.

[6]Notice that we assume the parties remain under non-integration after a deviation (as in Halonen (2002)), rather than reverting to the optimal static governance (integration), because integration elicits the first best effort and thus no relational contract can be built on such a valuable fallback

For effort this relational contract to be an equilibrium, asking A to exert effort a, and incurring the corresponding wage cost $\frac{a^2}{2}$, must be incentive compatible for R, that is:

$$-\frac{a^2}{2} + b + \frac{\delta}{1-\delta}(s + b - \frac{a^2}{2}) \geq \frac{\delta}{1-\delta} u_R^{spot},$$

where $u_R^{spot} = \frac{f^2}{8}$ is R's per period payoff under static non-integration, as derived in Section 4.2.

Similarly, for the relational contract to be an equilibrium, paying the bonus b must be incentive compatible for M, that is:

$$-b + \frac{\delta}{1-\delta}(fa - s - b) \geq \frac{\delta}{1-\delta} u_M^{spot},$$

where $u_M^{spot} = \frac{f^2}{4}$ is M's payoff under static non-integration.

A necessary condition for the relational contract to be self-enforcing is therefore that the sum of M's and R's incentive constraints be satisfied:

$$\frac{\delta}{1-\delta}(fa - \frac{a^2}{2} - V^{spot}) \geq \frac{a^2}{2},$$

where $V^{spot} = u_R^{spot} + u_M^{spot} = \frac{3f^2}{8}$ is M's and R's per period joint surplus under static non-integration.

So long as this condition is satisfied, a contract with $s = u_R^{spot}$ and $b = \frac{a^2}{2}$ satisfies both M's and R's individual incentive constraints at the minimum cost for M. Thus, the joint incentive constraint is also sufficient for our relational contract to be an

option (Baker et al., 2002). Reverting to integration may indeed not be feasible in our empirical setting due to the existence of long-term outsourcing contracts. Additionally, one could easily reinterpret our model as a reduced form for a richer model in which integration achieves the first best effort but reduces efficiency on some other dimension, such as cost efficiency. Such a model would generate results that are qualitatively similar to those shown here.

equilibrium. M therefore chooses effort in the relational contract to solve

$$max_a \quad \{fa - \frac{a^2}{2} - u_R^{spot}\},$$

subject to the joint incentive constraint above. It immediately follows from the joint incentive constraint that this relational contract elicits the first best effort $a^{FB} = f$ if the constraint is slack, that is, if $\delta \geq 0.8$. If instead $\delta < 0.8$, the incentive constraint is binding and effort is given by the value of a that solves the joint incentive constraint with equality or, if such value does not exist, by the effort under static non-integration. Either way, effort under $\delta < 0.8$ is below the first best level.[7]

In other words, the model predicts that relational contracts under non-integration will replicate the vertical integration outcome if M and R have a high-value relationship ($\delta \geq 0.8$), and will perform worse than vertical integration if the relationship between M and R is more arm's-length ($\delta < 0.8$).

In the next section we bring the predictions of our model to the data.

5 DATA AND MEASURES

To conduct our empirical study, we combined four different databases. The first one comes from the Official Airline Guide (OAG hereafter), which provides information on the flight's operating and ticketing airline, flight number, airport, day and scheduled landing hour for all flights landing at a US airport during the winter quarter of 2017. The second dataset is the On-Time Performance data from the BTS, which provides information on delays for all domestic US flights, except for flights operated by small domestic airlines, mostly major-fully-owned and independent regional

[7] It is easy to check that for $\delta < 0.75$, there is no real effort satisfying the joint incentive constraint with equality, implying that the only feasible outcome is the static non-integration effort $a^{NI} = \frac{f}{2}$. For $\delta \in [0.75, 0.8)$, the effort with relational contracts is $a^{RC} = \delta f + \frac{f}{2}\sqrt{\delta(4\delta - 3)} < f$.

airlines. We complement the publicly available information from the BTS with data from FlightAware, which contains delay information from flights excluded in the BTS dataset. Our fourth and last dataset reports all instances in which a flight is rescheduled to facilitate another flight's landing during airport-based GDPs mandated by the FAA. We obtained this information for all U.S. airports during February 2017 by submitting a Freedom of Information Act request to the FAA.[8]

We measure (the inverse of) a flight's on-time performance, *Delay*, as the difference between its arrival at the gate and the scheduled arrival. We then use the OAG data to measure the governance form of each flight. Specifically, we construct a dummy for whether a flight is directly operated by the major, VI, and a dummy for whether the flight is operated by a regional partner that is fully owned by the major, VIP. Additionally, we use the methodology developed by Gil et al. (2022) to construct dummies for whether conditional on a flight being outsourced, the relationship between the major and the regional partner operating that flight has high or low present discounted value (PDV).

The key idea underlying this methodology is that in order to secure the participation of their regional partners' in slot exchanges, the majors enter relational contracts with them, which are sustained by the PDV of their ongoing relationship. When designing their network, the majors prefer to outsource the worse-weather routes to their high-PDV partners because on those routes, the partner is asked more often to participate in slot exchanges and thus has a higher temptation to renege on the slot exchange promise. To ensure the promise is self-enforcing despite a high reneging temptation, a high PDV of the relationship is needed, implying that aggregate

[8]More details on the construction of the FAA database and on slot exchanges can be found in Argyres et al. (2025), which uses a subset of this data (on the exchanges between majors and outsourced partners) to study how adaptation is influenced by a major's decision to use few versus multiple partners.

adverse weather across the routes a major outsources to a given partner is a lower bound for the PDV of their relationship. Accordingly, we construct three alternative measures of high PDV, based on different definitions of adverse weather. For each route in our data, we estimate the maximum rain precipitation, snow precipitation and number of freezing months for both endpoints of the route and take only the maximum value of the two endpoints.[9] Once these values are created, we take the average weather for each of these three dimensions across all routes outsourced for each major-regional pair. Finally, we create three dummy variables that take value 1 if the value is in the top 75th percentile of the distribution of values across all 38 major-regional relationships in our data. In a sense, these dummies take value 0 when the major-regional relationships have the lowest PDV value according to our measure in Gil et al. (2022).[10]

We also construct dummies for whether a flight occurs in a day in which a GDP is enacted by the FAA, GDP, and a second dummy, SE, for whether conditional on a GDP being in place, the flight is included in a slot exchange package. Lastly, we construct several control variables to be included in our regressions. Specifically, we measure a flight's number of seats ($Seats$) and ticket price ($FareperMile$), and we construct dummies for both the general aircraft type and the specific aircraft used to operate the flight. We present summary statistics for all variables in Table 1.

[TABLE 1 HERE]

Table 1 shows summary statistics for our full sample of 627822 flights (top panel) as well as subsamples for flights occurring in airports during GDPs (mid panel) and those flights included in slot exchanges (bottom panel). It is easy to see that the

[9] To do so, we use NOAA data at the weather-station level from 1940 to 1960. We match the location of each station to each airport in our data and then proceed to create yearly averages and max values per year.

[10] More details on the construction of the PDV measures and on the underlying theory can be found in Gil et al. (2022).

average delay in the data is 5 minutes, but these increase to 41 minutes during GDPs and those flights involved in slot exchanges accrue delays of 77 minutes on average. Note there are no differences across episodes (GDP and SE) on the number of seats or fare per mile, which is telling of the exogeneity of the GDP occurrence and the widespread impact on flights included in slot exchanges.

6 RESULTS

6.1 Empirical Methodology

The goal of our paper is to study how a flight's governance (vertical integration and high- vs. low-value outsourcing relationships) affects its performance. Accordingly, Our first exercise is to study how shocks to the need for adaptation (GDPs) affect the on-time performance of flights under different flight governance forms. To do so, ein our baseline exercises we estimate by OLS linear regression models of the following type:

$$Delay_{f,m,r,a,d} = \alpha + \beta_1 \cdot G_{f,m,r,a,d} + \gamma_1 \cdot GDP_{a,d} + \gamma_2 \cdot SE_{f,m,r,a,d} + \delta \cdot \mathbf{X}_{f,m,r,a,d} + \lambda_{m,r} + \varepsilon_{f,m,r,a,d} \tag{1}$$

$Delay_{f,m,r,a,d}$, the dependent variable, measures the delay (in minutes) with which flight f ticketed by major m, serving route r (directional connection between two airports) and landing at airport a on day d. $G_{f,m,r,a,d}$ is an indicator for whether the flight's governance (more precisely defined below). $GDP_{a,d}$ is an indicator for whether there is a GDP at airport a on day d. $SE_{f,m,r,a,d}$ is an indicator for whether the flight is part of a slot exchange package submitted by the major. $\mathbf{X}_{f,m,r,a,d}$ is a vector of independent variables, including the controls and interactions between independent variables and the governance dummies. The error term $\varepsilon_{f,m,r,a,d}$ is assumed to be normally distributed and iid.

To ensure that our estimates of the effect of governance on performance are not driven by the major's control over slots at the airport endogenous selection of flights into governance forms, and to control for major airline capabilities and route characteristics (such as weather, airport congestion and passenger traffic) that may affect flight performance, we exploit the granularity of our data and include both major*route fixed effects ($\lambda_{m,r}$) in our regressions. Additionally, we include and day-of-month fixed effects and fixed effects for the flight's departure hour at the airport of origin to ensure that our results are not driven by the endogenous selection of flights scheduled on certain days or hours into specific governance forms, and to control for any features of February 2017 days that may affect traffic and airport congestion on and around a flight, such as holidays and weekday-specific patterns in our regressions. In some regressions we also include plane type fixed effects to control for the possibility that planes with different degrees of efficiency may be systematically used under different governance forms.

6.2 Vertical Integration and Performance Under Fixed Decision Control

In our first exercise, we estimate regression equation (1) using VI and VIP – our dummies for whether a flight is operated by the major or by a major-owned regional, respectively – as the governance variables. This regression therefore does not distinguish between outsourcing relationships with high or low PDV. We present our econometric estimations in Table 2 below.

[TABLE 2 HERE]

Two key insights emerge from our analysis. First, vertically integrated flights are systematically less delayed than outsourced flights. In days without a GDP, in which flights are operated according to routine procedures, the delay reduction in major-operated flights is about 6.6 minutes (column 4). Column 5 shows that the performance advantage of major-operated flights over outsourced flights goes up to 9 minutes when controlling for the number of seats (and assuming an aircraft of 65 seats, which is a size observed in both the smallest major-operated flights and the largest regional flights). On GDP days, the delay reduction in vertically integrated flights increases by an extra three minutes (columns 6-7), which is consistent with the fact that employees' efforts are more crucial at the margin when schedules are disrupted and there is more competition for airport resources.

Overall, these findings support the incentive/management view of firm boundaries illustrated by our model: holding the major's direct control over decisions affecting flight on-time-performance (namely, slots and airport resources), a flight's delay decreases when the agents responsible for it are employed by the major than when they are employed by an outsourced partner. It is interesting to observe that while flights operated by major-owned regional partners are also more timely than outsourced flights (the VIP coefficient is negative in all regressions), their performance advantage is much smaller than for major-operated flights. This finding is also consistent with the incentive/management view of integration: while the major has the same control of alienable decisions for major-operated flights and flights operated by a major-owned regional, the integrated regional's employees are not directly managed by the major, who may therefore suffer from some loss of managerial control, albeit lower than in the case of outsourced flights.

The control variables also have the expected effects on flight delays. Flights are more severely delayed on GDP days and when they are involved in a slot exchange. Moreover, larger flights experience lower delays, possibly due to the fact that they are more valuable to the major and hence are allocated more resources and effort.[11]

In Tables 3 and 4 below we add plane fixed effects to the above regressions. Table 3 includes general aircraft type fixed effects whereas Table 4 use specific aircraft model fixed effects. Controlling for plane type holds constant any plane features (beyond the number of seats) that may simultaneously affect the major's choice of operating airline and the plane's on-time-performance. The results are entirely consistent with those in Table 2.

[TABLES 3 AND 4 HERE]

Notice that when using aircraft model fixed effects, it is no longer possible to identify the VI coefficient because the planes used by majors are systematically different from the regional jets used by outsourced regional partners, and thus there is no variation in the VI variable within an aircraft model. However, major-owned and outsourced regional partners both use regional aircraft models, and thus it is possible in Table 4 to estimate the effect of VIP coefficient, which is negative and statistically significant in most specifications.[12]

6.3 Vertical Integration Versus Relational Contracts

In our second exercise, we add to equation (1) our three proxies for whether an outsourced flight is operated by a regional that has a high-value relationship with the major. We present our econometric estimations in Tables 5 and 6 below.

[TABLES 5-6 HERE]

[11] Appendix Table A1 shows that our results are robust to restricting attention to the sample of flights operated by major airlines only, namely, American Airlines, Delta Airlines, United Airlines and Alaska Airlines. These are the only US airlines outsourcing flights to independent regional airlines.

[12] Appendix Table A2 uses Fare per mile as potential proxy for route value.

Table 5 estimates the effect of high-PDV outsourcing relationships on delays, and compares it to the effect of vertical integration, using the full sample of flights. Table 6 does the same but restricts attention to the subsample of days in which no GDP is in place. We break the sample rather than interacting the PDV proxies with the GDP indicator, unlike Table 2 through 5, to facilitate table readability given the large number of independent variables and interactions.

The results indicate that flights operated by an outsourced partner whose relationship with the major has high PDV (as proxied by the major's choice to assign adverse-weather routes to that partner) experience significantly lower delays than flights outsourced to low-PDV partners: the coefficients of the high-precipitation, high-snow and low-freezing-months network indicators are negative and significant in all specifications. To illustrate, consider column 2 in Table 6: on a non-GDP day, flights outsourced to a high-PDV partner, as proxied by the network precipitation dummy, experience delays that are seven minutes lower than the delays of flights outsourced to low-PDV partners. The delay advantage of high-PDV outsourced flights over low-PDV ones is therefore similar in size to the advantage of major-operated flights, which is also seven minutes. Similar patterns arise in Table 5, although the interpretation of coefficients there is more complicated because the PDV effects may be partially driven by flights that are selected into slot exchange packages during GDP days.

These results are overall consistent with Gibbons's observation that comparing vertical integration to a bundle of arm's length and relational outsourcing, as most of the literature does, may systematically bias our assessment of the effect of vertical integration on performance (Gibbons, 2025). Tables 5 and 6 suggest that the performance advantage of vertical integration is driven by low-value outsourcing relationships in which relational contracts between the major and the regional partners are harder to sustain. In high-value relationships, relational contracts appear to substitute well for vertical integration. This is an important observation because vertical integration may be costly on several dimensions (outside of the scope of this study), and thus it is important for managers to know that when relational contracts are feasible, there is no big loss in using them (instead of vertical integration) as a tool to govern suppliers.

REFERENCES

Acemoglu, D., Aghion, P., Griffith, R., and Zilibotti, F. (2010). Vertical integration and technology: Theory and evidence. *Journal of the European Economic Association*, 8:989–1033.

Argyres, N., Gil, R., and Zanarone, G. (2025). Does supply concentration encourage cooperation? evidence from airlines. *Management Science*, 71:5083–5105.

Baker, G., Gibbons, R., and Murphy, K. (1994). Subjective performance measures in optimal incentive contracts. *Quarterly Journal of Economics*, 109(4):1125–1156.

Baker, G., Gibbons, R., and Murphy, K. (2002). Relational contracts and the theory of the firm. *Quarterly Journal of Economics*, 117:39–84.

Baker, G. and Hubbard, T. (2004). Contractibility and asset ownership: On-board computers and governance in u.s. trucking. *Quarterly Journal of Economics*, 119(4):1443–1479.

Barron, D., Gibbons, R., Gil, R., and Murphy, K. (2020). Relational adaptation under reel authority. *Management Science*, 66:1868–1889.

Bolton, P. and Dewatripont, M. (2004). *Contract Theory*. MIT Press.

Coase, R. (1937). The nature of the firm. *Economica*, 4:386–405.

Dessein, W. (2014). Incomplete contracts and firm boundaries: New directions. *Journal of Law, Economics and Organization*, 30:i13–i36.

Forbes, S. and Lederman, M. (2009). Adaptation and vertical integration in the airline industry. *American Economic Review*, 99(5):1831–1849.

Forbes, S. and Lederman, M. (2010). Does vertical integration affect firm performance? evidence from the airline industry. *RAND Journal of Economics*, 41(4):764–790.

Gibbons, R. (2005a). Four formalizable theories of the firm. *Journal of Economic Behavior and Organization*, 58:200–245.

Gibbons, R. (2005b). Incentives Between Firms (and Within). *Management Science*, 51:2–17.

Gibbons, R. (2025). From coase to culture? visible hands build equilibria. *Economics*, forthcoming.

Gil, R., Kim, M., and Zanarone, G. (2022). Relationships under stress: Relational outsourcing in the u.s. airline industry after the 2008 financial crisis. *Management Science*, 68(2):1256–1277.

Gil, R. and Zanarone, G. (2017). Formal and informal contracting: Theory and evidence. *Annual Review Law and Social Sciences*, 13:141—-159.

Gil, R. and Zanarone, G. (2018). On the determinants and consequences of informal contracting. *Journal of Economics and Management Strategy*, 27:726—-741.

Grossman, S. and Hart, O. (1986). The costs and benefits of ownership: A theory of vertical and lateral integration. *Journal of Political Economy*, 94:691–719.

Halonen, M. (2002). Reputation and the allocation of ownership. *Economic Journal*, 112:539–558.

Hansman, C., Hjort, J., Leon-Cillotta, G., and Teachout, M. (2020). Vertical integration, supplier behavior and quality upgrading among exporters. *Journal of Political Economy*, 128(9):3570–3625.

Hart, O. and Moore, J. (1990). Property rights and the nature of the firm. *Journal of Political Economy*, 98:1119–1158.

Holmstrom, B. (1999). The firm as a subeconomy. *Journal of Law, Economics & Organization*, 15:74–102.

Holmstrom, B. and Milgrom, P. (1987). Aggregation and Linearity in the Provision of Intertemporal Incentives. *Econometrica*, 55:303–328.

Holmstrom, B. and Milgrom, P. (1991). Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design. *Journal of Law, Economics & Organization*, 7:24.

Innes, R. (1990). Limited Liability and Incentive Contracting with Ex Ante Action Choices. *Journal of Economic Theory*, 52:47–65.

Lafontaine, F. and Slade, M. (2007). Vertical integration and firm boundaries: The evidence. *Journal of Economic Literature*, 45:629–685.

Lafontaine, F. and Slade, M. (2013). Inter-firm contracts: Evidence. *Gibbons, R and Roberts, J (eds.), Handbook of Organizational Economics*, pages 958–1013.

Levin, J. (2003). Relational incentive contracts. *American Economic Review*, 93(3):835–857.

Macchiavello, R. (2022). Relational contracts and development. *Annual Review of Economics*, 14:337–362.

Macchiavello, R. and Miquel-Florensa, P. (2019). Vertical integration and inter-firms relationships in the costa rica coffee chain. *Working Paper*.

Macchiavello, R. and Morjaria, A. (2015). The value of relationships: Evidence from a supply shock to kenyan rose exports. *American Economic Review*, 105:2911–2945.

Macchiavello, R. and Morjaria, A. (2021). Competition and relational contracts in the rwanda coffee chain. *Quarterly Journal of Economics*, 136:1089–1143.

Masten, S. (1984). The organization of production: Evidence from the aerospace industry. *Journal of Law and Economics*, 27:403–417.

Masten, S. (1988). A legal basis for the firm. *Journal of Law, Economics and Organization*, 4:181–198.

Monteverde, K. and Teece, D. (1982). Supplier switching costs and vertical integration in the automobile industry. *The Bell Journal of Economics*, pages 206–213.

Sappington, D. (1983). Limited liability contracts between principal and agent. *Journal of Economic Theory*, 29:1–21.

Williamson, O. (1971). The vertical integration of production: Market failure considerations. *American Economic Review*, 61:112–123.

Williamson, O. (1979). Transaction cost economics: The governance of contractual relations. *Journal of Law and Economics*, 22:233–261.

Williamson, O. (1991). Comparative economic organization: The analysis of discrete structural alternatives. *Administrative Science Quarterly*, 36:269–296.

TABLES

Table 1. Summary Statistics of Main Variables in Full Sample, GDP and SE subsamples

	No. Obs	Mean	Std. Dev.	Min	Q1	Q2	Q3	Max
Arrival Delay	627822	5.448	50.703	-1341.5	-16	-6	9.933	1854
VI (major-operated)	627822	0.656	0.475	0	0	1	1	0
VIP (operated by major-owned regional)	627822	0.093	0.291	0	0	0	0	1
Major	627822	0.727	0.445	0	0	1	1	1
GDP at airport	627822	0.141	0.348	0	0	0	0	1
SE (flight in slot exchange)	627822	0.068	0.251	0	0	0	0	1
Seats	627146	125.626	52.899	30	76	143	160	376
Fare per mile	551713	0.476	0.204	0.067	0.326	0.441	0.589	4.662
If GDP = 1								
Arrival Delay	88472	41.147	85.486	-1286	-8	15	65	1854
VI (major-operated)	88472	0.685	0.464	0	0	1	1	1
VIP (operated by major-owned regional)	88472	0.051	0.220	0	0	0	0	1
Major	88472	0.764	0.424	0	1	1	1	1
SE (flight in slot exchange)	88472	0.480	0.500	0	0	0	1	1
Seats	88445	130.622	53.705	37	76	143	166	374
Fare per mile	80,013	0.464	0.227	0.094	0.301	0.405	0.569	1.928
If SE = 1								
Arrival Delay	42489	77.334	98.028	-1285	17	53	114	1342
VI (major-operated)	42489	0.651	0.477	0	0	1	1	1
VIP (operated by major-owned regional)	42489	0.043	0.204	0	0	0	0	1
Major	42489	0.805	0.396	0	1	1	1	1
Seats	42483	125.637	51.402	50	76	143	160	366
Fare per mile	41272	0.481	0.226	0.104	0.317	0.430	0.576	1.928

This table provides summary statistics for all variables used in our analysis in the full working sample, sample of flights arriving to airports under GDP and a sample used in slot exchanges.

Table 2. Main Results with Full Working Sample

	1	2	3	4	5	6	7
Dependent Variable:	**Arrival Delay**						
VI (major-operated)	-7.295***	-6.115***	-32.189***	-6.681***	-34.026***	-4.803***	-33.335***
	(0.747)	(0.864)	(2.566)	(0.809)	(2.422)	(0.771)	(2.352)
VIP (operated by major-owned regional)	-5.038***	-2.196***	-19.946***	-1.955***	-21.226***	-1.003*	-22.130***
	(1.280)	(0.664)	(4.089)	(0.646)	(4.015)	(0.592)	(3.889)
GDP at airport				10.169***	10.158***	13.127***	13.276***
				(1.568)	(1.568)	(2.575)	(2.572)
SE (flight in slot exchange)				58.831***	58.865***	71.165***	71.354***
				(3.935)	(3.936)	(5.197)	(5.202)
VI*GDP						-3.711*	-3.909**
						(1.997)	(1.994)
VI*SE						-17.738***	-17.961***
						(3.986)	(3.983)
VIP*GDP						-3.775	-4.077
						(2.741)	(2.734)
VIP*SE						-12.78	-12.976
						(14.448)	(14.424)
Seats			-0.347***		-0.367***		-0.385***
			(0.030)		(0.029)		(0.029)
VI*Seats			0.362***		0.381***		0.399***
			(0.030)		(0.029)		(0.029)
VIP*Seats			0.274***		0.297***		0.326***
			(0.063)		(0.061)		(0.061)
Constant	10.706***	9.654***	32.796***	4.587***	29.077***	3.219***	28.900***
	(1.325)	(0.720)	(2.123)	(0.608)	(1.947)	(0.600)	(1.967)
FE Day of Month	NO	YES	YES	YES	YES	YES	YES
FE Major-Route	NO	YES	YES	YES	YES	YES	YES
FE departure hour-Origin Airport	NO	YES	YES	YES	YES	YES	YES
Observations	627,822	627,633	626,963	627,633	626,963	627,633	626,963
R-squared	0.004	0.176	0.176	0.234	0.235	0.236	0.236

Variable definitions

vi = dummy for whether flight operated by major; vip = dummy for whether flight operated by major-owned regional
(vi & vip are mutually exclusive dummies, the omitted reference category is "flight operated by outsourced regional")
GDPa = dummy for whether there is GDP at airport*day; GDPf = dummy for whether flight is in slot exchange package.
Robust standard errors clustered at the arrival airport/day of the month level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 3. Main Results with Full Working Sample and General Aircraft Type Fixed Effects

	1	2	3	4	5	6
Dependent Variable:	Arrival Delay					
VI (major-operated)	-6.204***	-20.092***	-6.241***	-22.110***	-3.355***	-19.332***
	(1.620)	(3.348)	(1.323)	(3.204)	(1.175)	(3.065)
VIP (operated by major-owned regional)	-4.258***	-17.239***	-4.092***	-18.304***	-3.258***	-18.224***
	(0.654)	(3.691)	(0.632)	(3.626)	(0.603)	(3.593)
GDP at airport			10.048***	10.043***	13.503***	13.472***
			(1.572)	(1.572)	(2.560)	(2.561)
SE (flight in slot exchange)			58.307***	58.325***	70.667***	70.725***
			(3.921)	(3.922)	(5.177)	(5.179)
VI*GDP					-4.370**	-4.326**
					(1.982)	(1.982)
VI*SE					-17.924***	-17.987***
					(3.941)	(3.939)
VIP*GDP					-4.421	-4.481*
					(2.699)	(2.700)
VIP*SE					-12.379	-12.365
					(14.352)	(14.355)
Seats		-0.224***		-0.248***		-0.249***
		(0.037)		(0.036)		(0.036)
VI*Seats		0.193***		0.219***		0.220***
		(0.038)		(0.037)		(0.037)
VIP*Seats		0.214***		0.235***		0.247***
		(0.057)		(0.055)		(0.057)
Constant	9.710***	26.843***	4.465***	23.231***	2.462***	21.292***
	(1.188)	(2.974)	(0.882)	(2.847)	(0.818)	(2.756)
FE Day of Month	YES	YES	YES	YES	YES	YES
FE Major-Route	YES	YES	YES	YES	YES	YES
FE departure hour-Origin Airport	YES	YES	YES	YES	YES	YES
FE General Aircraft Name	YES	YES	YES	YES	YES	YES
Observations	601,193	601,193	601,193	601,193	601,193	601,193
R-squared	0.17	0.17	0.23	0.23	0.23	0.23

Variable definitions

vi = dummy for whether flight operated by major; vip = dummy for whether flight operated by major-owned regional
(vi & vip are mutually exclusive dummies, the omitted reference category is "flight operated by outsourced regional")
GDPa = dummy for whether there is GDP at airport*day; GDPf = dummy for whether flight is in slot exchange package.
Robust standard errors clustered at the arrival airport/day of the month level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 4. Main Results with Full Working Sample and Specific Aircraft Type Fixed Effects

	1	2	3	4	5	6
Dependent Variable:	**Arrival Delay**					
VI (major-operated)	-	-	-	-	-	-
	-	-	-	-	-	-
VIP (operated by major-owned regional)	-3.878***	-20.508***	-3.619***	-21.583***	-2.895***	-21.077***
	(0.660)	(3.733)	(0.664)	(3.723)	(0.649)	(3.682)
GDP at airport			10.036***	10.031***	13.489***	13.479***
			(1.572)	(1.572)	(2.566)	(2.566)
SE (flight in slot exchange)			58.334***	58.349***	70.737***	70.774***
			(3.928)	(3.927)	(5.186)	(5.185)
VI*GDP					-4.367**	-4.353**
					(1.984)	(1.985)
VI*SE					-17.991***	-18.025***
					(3.951)	(3.950)
VIP*GDP					-4.418	-4.486*
					(2.694)	(2.698)
VIP*SE					-12.412	-12.411
					(14.338)	(14.340)
Seats		-0.198		-0.212		-0.235
		(0.185)		(0.172)		(0.163)
VI*Seats		0.128		0.130		0.152
		(0.186)		(0.173)		(0.164)
VIP*Seats		0.259***		0.281***		0.284***
		(0.058)		(0.056)		(0.057)
Constant	5.646***	17.338**	0.367	13.615*	0.248	14.025*
	(0.530)	(7.804)	(0.274)	(7.339)	(0.262)	(7.244)
FE Day of Month	YES	YES	YES	YES	YES	YES
FE Major-Route	YES	YES	YES	YES	YES	YES
FE departure hour-Origin Airport	YES	YES	YES	YES	YES	YES
FE Specific Aircraft Name	YES	YES	YES	YES	YES	YES
Observations	601,193	601,193	601,193	601,193	601,193	601,193
R-squared	0.17	0.17	0.23	0.23	0.23	0.23

Variable definitions
vi = dummy for whether flight operated by major; vip = dummy for whether flight operated by major-owned regional
(vi & vip are mutually exclusive dummies, the omitted reference category is "flight operated by outsourced regional")
GDPa = dummy for whether there is GDP at airport*day; GDPf = dummy for whether flight is in slot exchange package.
Robust standard errors clustered at the arrival airport/day of the month level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 5. Main Results with Outsourcing PDV proxies and Full Sample

	1	2	3	4	5	6	7
Dependent Variable:	**Arrival Delay**						
VI (major-operated)	-7.178***	-6.496***	-35.823***	-7.043***	-37.672***	-5.295***	-37.227***
	(0.996)	(1.012)	(2.640)	(0.918)	(2.547)	(0.921)	(2.467)
VIP (operated by major-owned regional)	1.357	0.734	-24.046***	0.655	-25.547***	1.235	-26.511***
	(1.095)	(0.929)	(4.115)	(0.881)	(4.040)	(0.869)	(3.868)
Dummy High Precipitation Outsourcing Network	-17.173***	-5.210**	-2.682	-4.572**	-1.935	-3.543	-0.780
	(3.154)	(2.332)	(2.461)	(2.254)	(2.384)	(2.297)	(2.428)
Dummy High Snow Outsourcing Network	-0.724	-3.205***	-4.495***	-3.008***	-4.359***	-3.550***	-4.972***
	(1.197)	(1.056)	(1.103)	(1.010)	(1.055)	(1.002)	(1.053)
Dummy High Freezing Months Network	16.255***	6.636***	8.744***	5.919***	8.141***	5.129***	7.453***
	(2.023)	(1.689)	(1.613)	(1.615)	(1.544)	(1.668)	(1.593)
GDP at airport				10.172***	10.161***	13.097***	13.214***
				(1.569)	(1.569)	(2.557)	(2.556)
SE (flight in slot exchange)				58.804***	58.842***	71.115***	71.382***
				(3.938)	(3.938)	(5.204)	(5.210)
VI*GDP						-3.709*	-3.886**
						(1.972)	(1.969)
VI*SE						-17.693***	-18.019***
						(3.986)	(3.977)
VIP*GDP						-3.334	-3.383
						(2.717)	(2.724)
VIP*SE						-12.855	-13.121
						(14.503)	(14.488)
Seats			-0.420***		-0.442***		-0.462***
			(0.038)		(0.037)		(0.037)
VI*Seats			0.435***		0.456***		0.477***
			(0.037)		(0.036)		(0.036)
VIP*Seats			0.369***		0.390***		0.414***
			(0.060)		(0.058)		(0.058)
Constant	10.588***	9.954***	36.344***	4.880***	32.636***	3.655***	32.679***
	(1.688)	(0.837)	(2.392)	(0.750)	(2.218)	(0.766)	(2.237)
FE Day of Month	NO	YES	YES	YES	YES	YES	YES
FE Major-Route	NO	YES	YES	YES	YES	YES	YES
FE departure hour-Origin Airport	NO	YES	YES	YES	YES	YES	YES
Observations	627,822	627,633	626,963	627,633	626,963	627,633	626,963
R-squared	0.01	0.18	0.18	0.24	0.24	0.24	0.24

Variable definitions

vi = dummy for whether flight operated by major; vip = dummy for whether flight operated by major-owned regional
(vi & vip are mutually exclusive dummies, the omitted reference category is "flight operated by outsourced regional")
GDPa = dummy for whether there is GDP at airport*day; GDPf = dummy for whether flight is in slot exchange package.
Dummies of High Precipitation, Snow and Freezing Months Outsourcing Network take value 1 if the average weather across routes outsourced by a major to a regional is above the 25th percentile across all 38 major-regional relationships in our data. The reference group in these regressions is then the set of relationships with the most favorable weather and, thus, according to our proxy, the lowest relationship values.
Robust standard errors clustered at the arrival airport/day of the month level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 6. Main Results with Outsourcing PDV proxies and non-GDP Sample

	1	2	3
Dependent Variable:	**Arrival Delay**		
VI (major-operated)	-3.166***	-7.307***	-35.728***
	(0.358)	(0.505)	(1.876)
VIP (operated by major-owned regional)	4.001***	2.044***	-24.615***
	(0.517)	(0.551)	(2.442)
Dummy High Precipitation Outsourcing Network	-8.902***	-7.440***	-4.990***
	(0.716)	(0.809)	(0.860)
Dummy High Snow Outsourcing Network	-3.826***	-1.848***	-3.043***
	(0.463)	(0.529)	(0.535)
Dummy High Freezing Months Network	13.838***	7.371***	9.251***
	(0.543)	(0.689)	(0.709)
Seats			-0.381***
			(0.027)
VI*Seats			0.409***
			(0.027)
VIP*Seats			0.398***
			(0.036)
Constant	1.151***	4.575***	28.061***
	(0.367)	(0.446)	(1.712)
FE Day of Month	NO	YES	YES
FE Major-Route	NO	YES	YES
FE departure hour-Origin Airport	NO	YES	YES
Observations	539,350	539,123	538,480
R-squared	0.01	0.10	0.10

Variable definitions

vi = dummy for whether flight operated by major; vip = dummy for whether flight operated by major-owned regional
(vi & vip are mutually exclusive dummies, the omitted reference category is "flight operated by outsourced regional")
Dummies of High Precipitation, Snow and Freezing Months Outsourcing Network take value 1 if the average weather across routes outsourced by a major to a regional is above the 25th percentile across all 38 major-regional relationships in our data. The reference group in these regressions is then the set of relationships with the most favorable weather and, thus, according to our proxy, the lowest relationship values.
Robust standard errors clustered at the arrival airport/day of the month level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A1. Main Results with Sample restricted to Flights with Major Ticketing Airline

	1	2	3	4	5	6	7
Dependent Variable:	**Arrival Delay**						
VI (major-operated)	-8.195***	-6.245***	-33.796***	-6.892***	-36.040***	-5.137***	-35.438***
	(0.740)	(0.782)	(2.700)	(0.730)	(2.591)	(0.703)	(2.535)
VIP (operated by major-owned regional)	-5.038***	-2.341***	-20.979***	-2.053***	-22.667***	-1.109*	-23.474***
	(1.280)	(0.676)	(3.660)	(0.646)	(3.601)	(0.595)	(3.513)
GDP at airport				10.229***	10.218***	13.382***	13.532***
				(1.788)	(1.788)	(2.537)	(2.535)
SE (flight in slot exchange)				62.009***	62.058***	71.431***	71.646***
				(4.271)	(4.273)	(5.160)	(5.167)
VI*GDP						-4.531**	-4.751***
						(1.837)	(1.838)
VI*SE						-15.044***	-15.310***
						(4.053)	(4.047)
VIP*GDP						-4.065	-4.392
						(2.697)	(2.690)
VIP*SE						-12.888	-13.077
						(13.857)	(13.830)
Seats			-0.344***		-0.370***		-0.387***
			(0.029)		(0.028)		(0.028)
VI*Seats			0.371***		0.396***		0.414***
			(0.030)		(0.029)		(0.029)
VIP*Seats			0.288***		0.318***		0.346***
			(0.056)		(0.054)		(0.055)
Constant	10.707***	9.315***	31.802***	3.453***	27.701***	2.360***	27.736***
	(1.325)	(0.701)	(2.097)	(0.522)	(1.937)	(0.520)	(1.952)
FE Day of Month	NO	YES	YES	YES	YES	YES	YES
FE Major-Route	NO	YES	YES	YES	YES	YES	YES
FE departure hour-Origin Airport	NO	YES	YES	YES	YES	YES	YES
Observations	456,450	456,296	456,087	456,296	456,087	456,296	456,087
R-squared	0.01	0.18	0.18	0.24	0.24	0.24	0.24

Variable definitions

vi = dummy for whether flight operated by major; vip = dummy for whether flight operated by major-owned regional
(vi & vip are mutually exclusive dummies, the omitted reference category is "flight operated by outsourced regional")
GDPa = dummy for whether there is GDP at airport*day; GDPf = dummy for whether flight is in slot exchange package.
Robust standard errors clustered at the arrival airport/day of the month level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A2. Main Results with Fare per mile Interactions

	1	2	3	4	5	6
Dependent Variable:	Arrival Delay					
VI (major-operated)	-6.885***	-3.996*	-7.034***	-5.797**	-3.969	-2.990
	(0.769)	(2.363)	(2.703)	(2.436)	(2.430)	(2.541)
VIP (operated by major-owned regional)	-5.059***	-1.597	6.087**	5.573**	6.987***	7.531***
	(1.343)	(3.155)	(2.856)	(2.562)	(2.697)	(2.321)
Fare per mile		5.273				
		(4.299)				
VI*Fare per mile		-4.660	1.655	-1.495	-1.242	-3.331
		(3.653)	(4.141)	(3.641)	(3.776)	(3.874)
VIP*Fare per mile		-5.796	-12.786***	-11.655***	-12.245***	-13.120***
		(3.886)	(4.193)	(3.906)	(4.015)	(3.441)
GDP at airport				9.968***	13.257***	13.423***
				(1.653)	(2.708)	(2.731)
SE (flight in slot exchange)				59.003***	71.301***	71.325***
				(3.956)	(5.270)	(5.277)
VI*GDP					-4.078*	-7.087**
					(2.085)	(3.104)
VI*GDP*Fare per mile						7.518
						(4.718)
VI*SE					-17.572***	-18.381***
					(4.127)	(6.233)
VI*SE*Fare per mile						1.55434275
						(9.037)
VIP*GDP					-4.895*	-7.565
					(2.745)	(5.468)
VIP*GDP*Fare per mile						4.001
						(8.152)
VIP*SE					-12.213	-16.760
					(14.939)	(26.494)
VIP*SE*Fare per mile						6.846
						(43.243)
Constant	10.986***	7.850***	10.325***	4.468***	3.074***	2.979***
	(1.401)	(3.017)	(0.915)	(0.772)	(0.747)	(0.769)
FE Day of Month	NO	NO	YES	YES	YES	YES
FE Major-Route	NO	NO	YES	YES	YES	YES
FE departure hour-Origin Airport	NO	NO	YES	YES	YES	YES
Observations	551,713	551,713	551,566	551,566	551,566	551,566
R-squared	0.003	0.003	0.18	0.24	0.25	0.25

Variable definitions

vi = dummy for whether flight operated by major; vip = dummy for whether flight operated by major-owned regional
(vi & vip are mutually exclusive dummies, the omitted reference category is "flight operated by outsourced regional")
GDPa = dummy for whether there is GDP at airport*day; GDPf = dummy for whether flight is in slot exchange package.
Robust standard errors clustered at the arrival airport/day of the month level in parentheses. *** p<0.01, ** p<0.05, * p<0.1