Simple maternal effect animal models may provide biased estimates of additive genetic and maternal variation

Joel L. Pick[1*], Craig A. Walling[1] & Loeske E. B. Kruuk[1]

[1] Institute of Ecology and Evolution, University of Edinburgh, Charlotte Auerbach Road, Edinburgh, EH9 3FL, UK

[*] Corresponding Author: joel.l.pick@gmail.com

ORCIDs

JLP: https://orcid.org/0000-0002-6295-3742

CAW: https://orcid.org/0000-0002-8547-9828

LEBK: https://orcid.org/0000-0001-8588-1123

Simple maternal effect animal models may provide biased estimates of additive genetic and maternal variation

Abstract

Maternal effects (the consistent effect of a mother on her offspring) can inflate estimates of additive genetic variation (V_A) if not properly accounted for. As they are typically assumed to cause similarities only among maternal siblings, they are often accounted for by modelling maternal identity effects. However, if maternal effects have a genetic basis, they create additional similarities among relatives with related mothers that are not captured by maternal identity effects. Unmodelled maternal genetic variance (V_{Mg}) may therefore still inflate V_A estimates in common quantitative genetic models, which is underappreciated in the literature. Using published data and simulations, we explore the extent of this problem. Published estimates from eight species suggest that a large proportion of total maternal variation (V_M) is genetic ($\sim$65%). Both these data and simulations confirmed that unmodelled V_{Mg} can cause overestimation of V_A and underestimation of V_M, the bias increasing with the proportion of non-sibling maternal relatives in a pedigree. Simulations show these biases are further influenced by the size and direction of any direct-maternal genetic covariance. The estimation of total additive genetic variation (V_{A_t}; the weighted sum of V_A and V_{Mg}) is additionally affected, limiting inferences about evolutionary potential from simple maternal effect models. Unbiased estimates require modelling V_{Mg} explicitly, but these models are often avoided due to perceived data limitations. We demonstrate that estimating V_{Mg} is possible even with small pedigrees, reducing bias in V_A estimates and maintaining accuracy in estimates of V_A, V_M, and V_{A_t}. We therefore advocate for the broader use of these models.

Keywords: Animal model; genetic variation; maternal effects; bias; wild population

Introduction

Mothers (and, more generally, parents) can have strong effects on the phenotype of their offspring, above the effect of their shared genes. These effects of the maternally provided environment on offspring phenotype are generally known as 'maternal effects' and, in quantitative genetics, refer specifically to the consistent effect of a mother across all her offspring (also known as 'maternal performance'; Riska *et al.*, 1985). Maternal effects are ubiquitous, occurring in a wide range of taxa (Roach & Wulff, 1987; Bernardo, 1996; Mousseau & Fox, 1998; Moore *et al.*, 2019b), with the strongest impact on traits expressed in juveniles (Wilson & Reale, 2006; Pick *et al.*, 2016a; Moore *et al.*, 2019b; Gauzere *et al.*, 2020b).

From an evolutionary perspective, maternal effects can impact the evolutionary potential of traits. Generally, we think of the response of a trait to selection (R) to be dependent on the additive genetic variation (V_A) in that trait and the strength of selection, summarised by the breeder's equation ($R = h^2S$; Lush, 1937) or Lande's gradient equation ($R = V_A\beta$; equation 7 in Lande, 1976, where h^2 is the heritability of the trait (V_A/V_P), S is the selection differential and β the selection gradient). Maternal effects have historically been seen as a 'troublesome' parameter in the estimation of selection response (Falconer, 1981), as not accounting for these shared effects across siblings can dramatically inflate estimates of V_A, and so lead to the overestimation of the evolutionary potential of a trait (e.g., Kruuk & Hadfield, 2007). However, when the traits mediating the maternal effects have a genetic basis (i.e., maternal *genetic* effects), then the response to selection of these traits is no longer predicted by V_A alone (even when maternal effects are properly modelled), and will instead be determined by the 'total' additive genetic variation available, which can be calculated as:

$$V_{A_t} = V_A + \frac{1}{2}V_{Mg} \tag{1}$$

(Dickerson, 1947; Willham, 1963, 1972), where V_{Mg} is the maternal genetic variance. When maternal genetic effects exist, V_{A_t} gives a better measure of evolutionary potential (Cheverud

& Moore, 1994; note that this assumes that selection only acts on the focal offspring trait and not directly on maternal performance (Cheverud, 1984; Hadfield, 2012), a point that we will return to in the Discussion). Maternal genetic effects represent a specific, yet ubiquitous, form of indirect genetic effects (IGEs), which more generally act to increase the total additive genetic variation (Moore *et al.*, 1997). To date, several empirical studies in wild populations have estimated a considerable genetic component to maternal effect variation (e.g. Wilson *et al.*, 2005b; Kruuk & Hadfield, 2007; McFarlane *et al.*, 2015), consistent with studies demonstrating genetic variation in parental care behaviours (Freeman-Gallant & Rothstein, 1999; Maccoll & Hatchwell, 2003; Walling *et al.*, 2008; Dor & Lotem, 2010; Adams *et al.*, 2015; Bell *et al.*, 2018; Räsänen & Kruuk, 2007).

Equation 1 assumes that direct and maternal genetic effects are independent. There may, however, be a genetic correlation between the direct genetic effects acting on an individual's trait and the maternal genetic effects that individual exerts on its offspring (for example, a shared genetic basis between juvenile size and parental provisioning). Indeed, studies from livestock show that direct and maternal genetic effects likely *negatively* covary (Wilson & Reale, 2006; Räsänen & Kruuk, 2007), although this correlation is probably small. Evolutionary potential is further dependent on such a joint genetic basis, with a negative genetic correlation lowering the evolutionary potential of a trait and vice versa. The evolutionary potential can therefore be fully described as:

$$V_{A_t} = V_A + \frac{3}{2}COV_{A,Mg} + \frac{1}{2}V_{Mg} \tag{2}$$

(Dickerson, 1947; Willham, 1963, 1972), with similar expressions being able to be made more broadly for IGEs (Moore *et al.*, 1997). Not incorporating maternal genetic variance and any direct-maternal genetic covariance (or IGEs more generally) will therefore further bias the estimation of the total evolutionary potential of a trait (Lynch, 1987; Kirkpatrick & Lande, 1989; Lande & Kirkpatrick, 1990; Wolf *et al.*, 1998; McGlothlin *et al.*, 2010).

Table 1: Glossary of symbols and terms.

Term	Definition
V_A	Additive genetic variance; the variation in direct additive genetic effects, which are also known as breeding values.
V_M	Total maternal variance ($V_{Mg} + V_{Me}$); the variation in a given phenotype due to the consistent effect of the environment that the individuals' mothers provide.
V_{Mg}	Maternal genetic variance; the part of the maternal variance that is due to genetic variation in maternal phenotypes.
V_{Me}	Maternal environmental variance; the part of the maternal variance that is due to environmental variation in maternal phenotypes. This is estimated (as $\hat{V}_{Me}$) by the maternal identity term in a full maternal effects model.
$COV_{A,Mg}$	Direct-maternal genetic covariance. This is the genetic covariance between an individual's direct genetic effect and the same individual's maternal genetic effect on its offspring (e.g., the genetic covariance between juvenile size and parental provisioning).
$\hat{V}_{Mc}$	Maternal identity variance; the variance estimated by a maternal identity term in a simple maternal effects model.
V_{A_t}	Total additive genetic variance, see equation 2. This is a measure of evolutionary potential, although it ignores that selection may act separately and even in opposite directions on offspring and maternal phenotypes.
Non-sibling maternal links	Links in a pedigree where the two individuals' mothers are related, that are not maternal siblings (e.g., mother-offspring; see Table 2). Calculation of the proportion of non-sibling maternal links for the purpose of this study involved only the relationships in Table 2.

Quantitative genetic studies in the wild typically focus on long-term studies of vertebrates, which show a considerable amount of maternal (or more generally parental) care. There is huge potential for maternal (genetic) effects in these systems, and consequently for V_A estimates (which we denote as $\hat{V}_A$) to be inflated when these effects are unaccounted for. Whereas these confounding effects can be accounted for using breeding designs in captive populations (Riska *et al.*, 1985; Cheverud & Moore, 1994; Lynch & Walsh, 1998), typically this has to be done statistically in wild populations (although note the use of cross fostering for this purpose; Cheverud & Moore, 1994). Animal models are currently the most common method for estimating additive genetic variation in the wild (Postma, 2014; Young & Postma, 2023a). They are an extension of a mixed model that allows the incorporation of relatedness information from the pedigree to estimate V_A (Henderson, 1988; Kruuk, 2004). Perhaps most importantly, they can be used to account for other sources of confounding variation, including a common environment (e.g. the maternal environment; Kruuk, 2004; Kruuk & Hadfield, 2007). They therefore allow for the explicit estimation of additive genetic and maternal effects. Although many studies of genetic variation in the wild have estimated maternal variance (Moore *et al.* (2019b) collated 770 estimates from 116 studies in the wild), by far the majority of these (97.8% of the estimates in Moore *et al.*, 2019b) do not estimate maternal *genetic* effects. Typically, maternal variance is estimated by including maternal identity as a random term in an animal model. We refer to these models as 'simple' maternal effects models, and the estimate of maternal variance from these models as $\hat{V}_{Mc}$. Although the paucity of 'full' maternal effects models (i.e. those estimating V_{Mg} and $COV_{A,Mg}$) may mainly be driven by the perception of data constraints (estimating maternal genetic variance requires more data, over more generations, which is often limited in studies of wild populations), we believe there is also a common assumption that all maternal genetic and maternal environment variance is captured by the maternal identity variance (i.e. $\hat{V}_{Mc} = V_M = V_{Mg} + V_{Me}$), in the same way that modelling individual identity captures permanent environment and additive genetic effects. Moore *et al.* (2019b), for example, sum V_{Me} and V_{Mg} estimates (which we denote as $\hat{V}_{Me}$ and $\hat{V}_{Mg}$) to get an estimate of total maternal variance ($\hat{V}_M$) to compare with models

that modelled only $\hat{V}_{Mc}$.

Although the assumption that $\hat{V}_{Mc}$ captures all V_{Mg} intuitively seems reasonable, several studies contain evidence that not directly modelling $\hat{V}_{Mg}$ and $C\hat{O}V_{A,Mg}$ may bias $\hat{V}_A$ even when maternal identity effects are modelled (simulations: Clément *et al.* 2001; Satoh *et al.* 2002; wild empirical studies: Wilson *et al.* 2005b; Kruuk & Hadfield 2007). We can see why this might happen when considering the sources of phenotypic covariance between different individuals in a pedigree, which are commonly presented in classic maternal effect papers from the animal breeding literature (Willham, 1963, 1972; Thompson, 1976, see Table 2). Whilst only maternal siblings (full siblings and maternal half-siblings) share maternal environmental effects (i.e., are raised by the same mother), any relatives whose mothers are related will share some V_{Mg}. Imagine two cousins related through their mothers being full siblings: as their mothers are related, the maternal genetic effects that the cousins experience will be similar due to their mothers' relatedness, but they will not share maternal *environment* effects (Figure 1A). In this way, the phenotypic covariance between any two individuals related via their mothers will include some degree of maternal genetic variation. On the other hand, two cousins with full-sibling fathers have unrelated mothers, and so will not share maternal effects of any form (Figure 1B). $COV_{A,Mg}$ is shared even more widely, by any two individuals who are related via one of their mothers. This means that even the phenotypic resemblance between fathers and their offspring includes some $COV_{A,Mg}$, as they are both related to the paternal grandmother (see the individual-sire covariance in Table 2). This is because the father's phenotype is affected by its inherited breeding value and the maternal genetic effect of its mother, which are correlated if there is $COV_{A,Mg}$ (Figure 1C). A positive correlation between these two leads to an increased resemblance between a father and his offspring, because the correlated maternal effect causes the father's phenotype to be even more like the offspring's breeding value than expected by the father's breeding value alone (and vice versa). Table 2 shows how different sources of variance contribute to the covariance between an individual and a variety of different relatives. These derivations are explained well in Lynch & Walsh

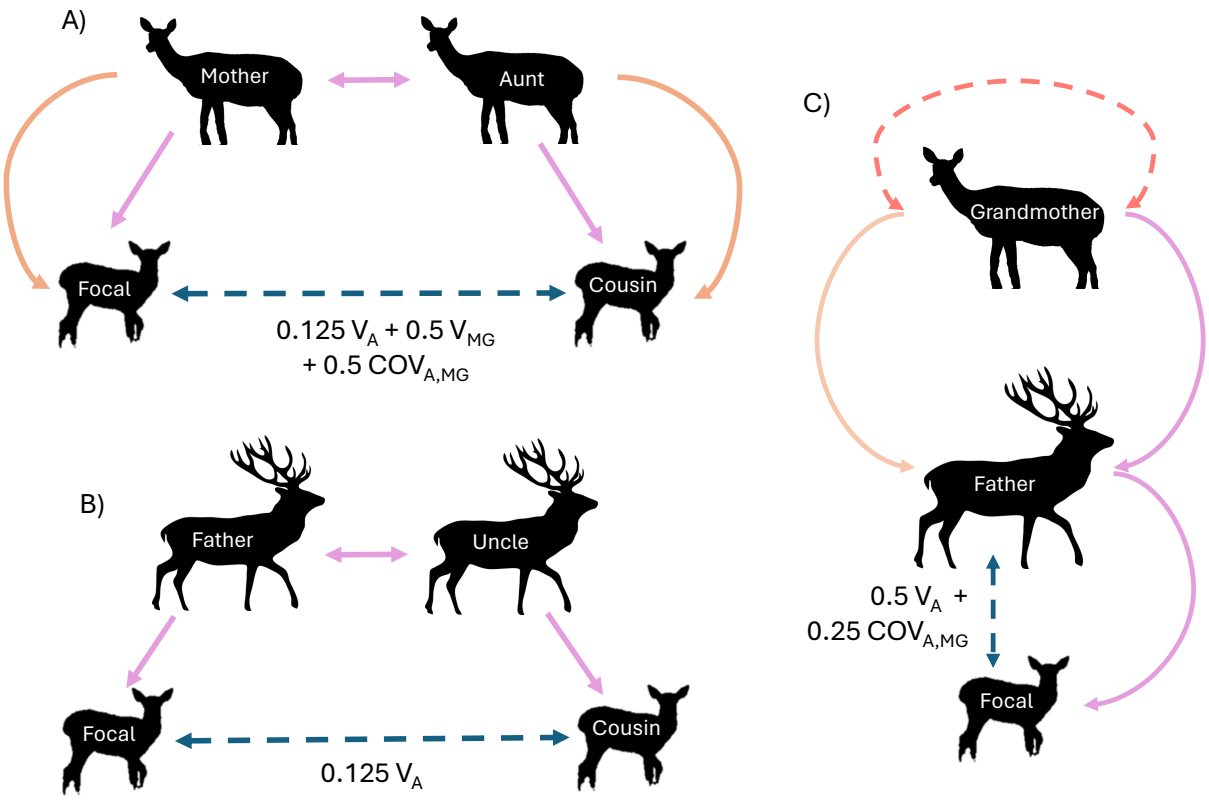


Figure 1: In the presence of maternal genetic effects, two cousin that are related via their mothers (A) will be more similar to each other than two cousins related via their fathers (B). Direct-maternal genetic covariance ($COV_{A,Mg}$) affects the phenotypic covariance between any individuals related via one or both of their mothers, for example fathers and their offspring (C), as the correlated maternal genetic effect causes the two individuals to be more or less similar to each other (depending on whether the covariance is positive or negative) than expected from the inherited breeding values alone. Orange arrows represent the maternal genetic effect, pink arrows the direct genetic effect (or relatedness) and the red dotted line the covariance between the two. The blue dotted line represents the phenotypic covariance between the two individuals of interest.

(1998, Chapter 23) and the full workings for Table 2 are show in the Supplementary Materials S1.

For multiple types of relationships, the presence of V_{Mg} and $COV_{A,Mg}$ generates covariance between relatives who do not share a mother and therefore do not share V_{Me}. Modelling $\hat{V}_{Mc}$ will, therefore, account for the maternal variance (environmental and genetic) shared between maternal siblings, but not the maternal genetic variance and covariance shared by other individuals related via their mothers. Consequently, we may expect that some of this

Table 2: The different components making up the expected phenotypic covariance between a focal individual and a given relative, for a set of close relationships. For example, the expected phenotypic covariance between an individual and its mother is $0.5V_A + 0.5V_{Mg} + 1.25COV_{A,Mg}$. In the aunt/uncle relationships, MHS and PHS refer to the relationships between the focal individual's parent and their sibling, indicating whether the parent and their sibling are maternal or paternal half sibs, respectively. These phenotypic covariances assume no inbreeding (i.e. an individual's parents are unrelated). Bold rows show those with non-sibling maternal links.

Relationship	Variances included in covariance			
	V_A	V_{Mg}	$COV_{A,Mg}$	V_{Me}
Dam	**0.5**	**0.5**	**1.25**	**0**
Sire	0.5	0	0.25	0
Full sib	0.5	1	1	1
Maternal half sib	0.25	1	1	1
Paternal half sib	0.25	0	0	0
Maternal grandparent	**0.25**	**0.25**	**0.625**	**0**
Paternal grandparent	0.25	0	0.125	0
Maternal full uncle/aunt	**0.25**	**0.5**	**0.75**	**0**
Paternal full uncle/aunt	0.25	0	0.25	0
Maternal half uncle/aunt (MHS)	**0.125**	**0.25**	**0.5**	**0**
Maternal half uncle/aunt (PHS)	0.125	0	0.25	0
Paternal half uncle/aunt (MHS)	0.125	0	0.25	0
Paternal half uncle/aunt (PHS)	0.125	0	0	0
Cousin - sires full sibs	0.125	0	0	0
Cousins - dams full sibs	**0.125**	**0.5**	**0.5**	**0**
Cousin - sire and dam full sibs	0.125	0	0.25	0
Cousin - sires half sibs	0.0625	0	0	0
Cousins - dams half sibs	**0.0625**	**0.25**	**0.25**	**0**
Cousin - sire and dam half sibs	0.0625	0	0.125	0

genetic variation that is not captured by $\hat{V}_{Mc}$ to be captured by $\hat{V}_A$. Unmodelled V_{Mg} would therefore be expected to bias $\hat{V}_A$ upwards (as this induces a positive covariance among maternal relatives), whilst $COV_{A,Mg}$ would bias $\hat{V}_A$ in the direction of the covariance (as it makes relatives resemble each other more when positive, and less when negative). The extent and direction of any bias in $\hat{V}_A$ from simple maternal effects models will therefore depend not only on the relative amounts of V_{Mg} and $COV_{A,Mg}$, but also on the structure of the pedigree, and in particular the degree to which the relationships in a pedigree are dominated by non-siblings that that share V_{Mg} and/or $COV_{A,Mg}$. We do not, however, know systematically to what extent estimates of $\hat{V}_A$ are affected, or the relative impact of pedigree structure. The only

simulation work to date (to our knowledge) focuses on breeding designs (and so pedigree structures) that are not very realistic to natural populations (Clément *et al.*, 2001; Satoh *et al.*, 2002). Pedigrees from natural populations will vary widely, due to factors such as life history variation, (sex specific) dispersal and mating system, with previous work showing that immigration has a large impact on the bias induced by unmodelled common environment effects (Kruuk & Hadfield, 2007). Given the wide usage of these simple maternal effect models in evolutionary ecology (Moore *et al.*, 2019b), it is important to understand the extent of the bias in the estimation of $\hat{V}_A$ caused by V_{Mg} and $COV_{A,Mg}$ if only $\hat{V}_{Mc}$ is modelled and how this may vary with pedigree structure. This will facilitate an understanding of how prevalent such biases may be in previous estimates of pedigree analyses from wild populations.

A final complication arises when we consider how these biases may affect the estimation of evolutionary potential. When using simple maternal effect models, the only measure of genetic variation comes from $\hat{V}_A$, meaning that we are implicitly assuming $\hat{V}_M$ is all environmental. If these models correctly estimated the underlying V_A and V_M and some or all of the maternal variance had a genetic basis, we would therefore be systematically underestimating evolutionary potential (measured as V_{A_t}; equation 2, the estimate of which is denoted as $\hat{V}_{A_t}$) by considering $\hat{V}_A$ alone, depending also on the direction and size of any direct-maternal genetic covariance (equation 2). On the other hand, any upward bias in $\hat{V}_A$ that occurs in the simple maternal effect models may compensate for this underestimation of evolutionary potential; even if $\hat{V}_A$ is biased, V_{A_t} may still be well estimated by $\hat{V}_{A_t}$. To our knowledge, only one empirical paper has compared $\hat{V}_{A_t}$ estimates from different models, finding that the estimates of total heritability of traits were similar between simple and complex maternal effect models (Table 2 in Wilson *et al.*, 2005b).

Here, to address the issues introduced above, we investigate several questions. First, we assess the extent to which maternal effects in wild populations have been shown to have a genetic component. Second, using simulations, we assess how simple maternal effect animal models are affected by the presence of unmodelled V_{Mg}, and to what extent these biases are affected

by pedigree structure. Third, we investigate the impact of $COV_{A,Mg}$ on these biases. Fourth, we assess the impact these biases have on our estimation of $\hat{V}_{A_t}$ and hence evolutionary potential. Finally, we investigate the feasibility of fitting full maternal effect models to small pedigrees as a means of mitigating these biases.

Methods

All simulations and analysis were carried out in R (version 4.2.1; R Core Team, 2022). In the reporting and description of our methods we follow the MeRIT guidelines (Nakagawa *et al.*, 2023).

Previous Estimates of $\hat{V}_{Mg}$

Although the extent of maternal genetic effects is well characterised in an animal breeding context (e.g., Wilson & Reale, 2006), the prevalence of these effects is much less well known in the wild. To this end, we made a non-exhaustive search for estimates of maternal genetic variation in wild populations using animal models. This search was not designed to be systematic, but we believe that it will have captured most estimates and so at least be representative. Initially, we were aware of four wild mammal species in which maternal genetic effects had been estimated, three of which Author 3 has been directly involved with. To find any additional estimates, Author 1 searched Web of Science on 22/05/2024 for the topic 'maternal genetic', subsetting by the Web of Science Categories 'Evolutionary Biology', 'Ecology' and 'Zoology', and for all papers citing the early wild maternal genetic effects papers (McAdam *et al.*, 2002; Wilson *et al.*, 2005b). All resulting papers were screened. Through this process, we discovered one further paper on a wild mammal species (Roe Deer; Quéméré *et al.*, 2018), two papers estimating such effects on different captive (but not domesticated) mammals (Blomquist & Williams, 2013; Ibáñez *et al.*, 2014) and one study on greenhouse plants (Galloway *et al.*, 2009). We excluded Gauzere *et al.* (2022) as it used a very similar dataset to Gauzere *et al.* (2020b). There are likely two sources of estimates that we may have

missed: papers that included these effects in their models, but they were not the main focus of the analysis, and unpublished studies. We have no reason to believe that the first source would be systematically different in size, and the second may be smaller due to publication bias. Our search gave 63 estimates of 8 species, from 14 studies (Soay sheep: Beraldi *et al.* 2007; Bérénos *et al.* 2014; Regan *et al.* 2017; Wilson *et al.* 2005b; Bighorn sheep: Wilson *et al.* 2005a; Réale *et al.* 2009; Red Deer: Kruuk & Hadfield 2007; Gauzere *et al.* 2020b, 2021; Roe Deer: Quéméré *et al.* 2018; Red Squirrel: McFarlane *et al.* 2015; Squirrel Monkey: Blomquist & Williams 2013; Culvier's Gazelle: Ibáñez *et al.* 2014; American Bellflower: Galloway *et al.* 2009). These estimates are shown in supplementary table S1.

From these studies Author 1 extracted the point estimates of $\hat{V}_A$, $\hat{V}_{Mg}$ and (where estimated) $\hat{V}_{Me}$. Most estimates were presented in tables, and we used the metaDigitise R package (version 1.0.1 Pick *et al.*, 2019a) to extract variance estimates from figures. We did not undertake a formal meta-analysis of these estimates - they came from few study systems, and in some of the systems included multiple estimates for the same traits from different papers. The model specifications were also different across studies. Furthermore, meta-analysis of variances estimated from mixed effects models is complicated for many reasons (such as zero-bounded estimates, presentation of frequentist point estimates and standard errors versus summaries of Bayesian posterior distributions, and inflation of effect sizes when power is low). Whilst we believe our presentation of these estimates gives a reasonable impression of the available estimates, these caveats should be borne in mind.

As maternal effects are likely to be stronger at earlier life stages, we categorised the estimates as being for phenotypic traits measured in the first year (n=34) vs at older ages (n=29). From these estimates we present the proportion of total phenotypic variation due to $\hat{V}_{Mg}$ ($\hat{m}^2_g$), and the proportion of total $\hat{V}_M$ due to additive genetic effects ($\frac{\hat{V}_{Mg}}{\hat{V}_{Mg}+\hat{V}_{Me}}$). To calculate the latter of these measures, we subset the data to only consider models that estimated both $\hat{V}_{Mg}$ and $\hat{V}_{Me}$ (in some cases only $\hat{V}_{Mg}$ was estimated), and those estimates where the total phenotype variation due to $\hat{V}_M$ ($\hat{m}^2$) was above 0.05, as the proportions were very unstable when $\hat{m}^2$ was

very low (the proportions were all exactly 0, 0.5 or 1; see Figure S3). As $\hat{V}_M$ represents the variation due to consistent effects of a mother on her offspring, this is similar to calculating $\frac{V_A}{V_A+V_{PE}}$, rather than a typical heritability (where V_{PE} is the 'permanent environment effects' variance; Kruuk & Hadfield, 2007). Additionally, from studies where different analyses were available (n=38), we took estimates of $\hat{h}^2$ and $\hat{m}^2$ from simple and complex maternal effects models and calculated the difference between them (simple model estimate - complex model estimate). We used estimates of the simple maternal model in McFarlane *et al.* (2014) to compare with some of the estimates in McFarlane *et al.* (2015). If the simple models typically overestimate V_A, then we expect the difference in $\hat{h}^2$ to be generally positive across these comparisons.

Pedigree Simulations

Maternal environment effects will only be shared by siblings, but maternal genetic effects will be shared by a larger set of maternal relatives, both siblings and non-siblings (Table 2). Therefore, whilst $\hat{V}_{Mc}$ estimates in simple maternal effects models will completely capture the V_{Me}, we expect that V_{Mg} will contribute to both $\hat{V}_{Mc}$ and $\hat{V}_A$ estimates, and that the bias in $\hat{V}_A$ may depend on the relative numbers of non-sibling maternal relatives. As a first step in testing this, we used individual-based simulations to generate pedigrees that would vary in their structure in a realistic way, and consequently vary in the proportion of non-sibling maternal links. Here, we varied three parameters in the pedigree simulations: the mean number of offspring per mother, the breeding system (monogamy versus polyandry) and sex-specific immigration rates.

The pedigrees were simulated by Author 1 using the pedAgree R package (version 0.0.1; Pick, 2024a). Across all pedigrees we simulated 5 discrete generations in addition to the founder population, a fixed population size (though the relative number of adults vs juveniles varied across pedigrees; see below), and a constant equal sex ratio. In all pedigrees 600 offspring were generated per generation, but the number of breeding females varied by pedigree type

according to the fecundity (see below). We assumed that all individuals with known parents (i.e., all individuals 'born' in the population) had a phenotype, meaning that all pedigrees had the same number of phenotyped individuals (3000), although varied slightly in the total size of the pedigree. These are large pedigrees compared to those typically studied in the wild; it is above the 95th quantile of sample sizes from studies using animal models based on the database from Young & Postma 2023a (using data from Young & Postma, 2023b). We deliberately used large pedigrees in these simulations to ensure that any bias is not due to low sample size.

We simulated three mating systems, defined by the resulting proportions of full vs half siblings within families: full-sib (each female only ever mates with one male and each male only ever mates with one female, so all offspring within a family are full sibs), intermediate (probability of 0.75 that a paired male sires the offspring, so families are a mixture of full and half sibs) and half sib (paternity of offspring was randomly assigned across all males, so offspring were almost always half siblings). Immigration was simulated as a certain proportion of breeders in each generation having unknown parents. We simulated four immigration scenarios: No immigration (closed population), unbiased immigration (25% of breeding females and 25% of breeding males were immigrants), female biased immigration (40% female, 10% male) and male biased immigration (10% female, 40% male). Finally, we simulated three fecundity scenarios (low, medium, and high). These were broadly based on the mean (lifetime) number of offspring per female from the 19 populations used in Bonnet *et al.* (2022), to ensure they were within realistic bounds for commonly studied wild animal populations. Low fecundity was three offspring per female, medium was six and high was 12. Pedigrees were simulated so that all females in the pedigree had the same fecundity. Sex-specific juvenile survival was dependent on the fecundity and immigration rates (2*(1 - immigration)/fecundity), to ensure a constant population size. We simulated 100 pedigrees for each of the 36 combinations of the mating system, immigration, and fecundity scenarios (3 x 4 x 3 = 36 combinations, 3600 pedigrees in total). Given the large sample size within each of the pedigrees, this number of

simulations is sufficient to well estimate bias.

For each of these pedigrees we calculated the proportion of non-sibling maternal links, by calculating the number of each of the non-sibling maternal relationships shown in Table 2 (shown in bold) as a proportion of the total number of the relationships in Table 2. Note, that the number of these relationships in a given pedigree can now be calculated using the pedtricks R package (version 0.5.0 onwards Martin *et al.*, 2024). Relationships were counted when both individuals had a phenotype (i.e., had a known mother). We were not seeking to create a metric to exactly predict the bias, but simply demonstrate the impact of the pedigree structure on the bias. The relationships shown in Table 2 are likely the most influential relationships, and so will largely capture the meaningful difference between pedigrees in the informative relationships. Figure S4 shows how the proportion of these links varies across pedigree types.

To check that the proportion of non-sibling maternal links in the simulated pedigrees were within a sensible range, we compared these values to two known pedigrees from real animal populations; a pedigree of a large ungulate (red deer, *Cervus elaphus*) used in Gauzere *et al.* (2020b) (data from Gauzere *et al.*, 2020a) and a hole nesting passerine bird (blue tit, *Cyanistes caeruleus*) used in Pick *et al.* (2022) (data from Pick & Hadfield, 2022). These pedigrees are broadly representative of many wild animal populations, and both have been used to analyse parental effects in juvenile size. It is important to note that metrics used to describe the pedigree are relative to the 'pruned' pedigree used in any given analysis (i.e., restricted to individuals relevant to analysis of a particular phenotypic trait), *not* the general characteristics of the whole population. Depending on what phenotypic trait is being analysed, and when in the life cycle it is expressed, the informative component of a pedigree will change, as it will be pruned for informative relationships prior to analysis (Morrissey & Wilson, 2010). For example, in a system with high numbers of offspring but low recruitment (such as blue tits), the structure of the pruned pedigree will change dramatically between analysing juvenile and adult traits, dependent on which individuals have phenotypes measured at which stages. Pedigrees

for juvenile traits will include most individuals born in the population, whereas pedigrees for adult traits will only comprise of those surviving to adulthood and immigrants, meaning the average number of offspring per mother will dramatically change. To demonstrate this, we generated a juvenile pruned pedigree and an adult pruned pedigree for each of the two full real pedigrees, by assuming that all individuals with known mother had phenotypes in the juvenile pedigree, and all individuals with offspring (i.e., recruits) had phenotypes in the adult pedigree. These values are shown in Figure 3.

Our simulated pedigrees are relatively deep (5 complete generations), and pedigree depth has previously been shown to affect pedigree structure and biases in V_A (Kruuk & Hadfield, 2007). In the Supplementary Materials (Section S4), we explored how pedigree depth affects the build-up of non-sibling maternal links and found that the proportion of non-sibling maternal links increases with pedigree depth, but stabilises after a couple of generations (Figure S6).

Simulated scenarios and models

We simulated a dataset from each of 12 scenarios with varying parameter sets (see Table 3) for each of the 3600 pedigrees, resulting in 43,200 datasets. Simulations were performed by Author 1 using the squidSim R package (version 0.2.3; Pick, 2024b). Phenotypes were simulated to have 0 mean and unit variance. Simulated V_A, V_{Me} and V_{Mg} therefore represented their respective proportions of total phenotypic variance explained (e.g., $V_A = h^2$). We assumed all effects were additive, and there was no dominance genetic variance (note the covariances shown in Table 2 also assume no dominance). We make this assumption for simplicity and because few studies in the wild model dominance effects and so implicitly assume no dominance variance (Ovaskainen *et al.*, 2008; Class & Brommer, 2020). V_A, V_{Me} and V_{Mg} were simulated to be either 0, 0.25 or 0.5, with varying genetic correlations (-0.6, -0.3, 0, 0.3 and 0.6), in different combinations to represent 12 different scenarios (shown in Table 3). We did not simulate all combinations of these values, but focused on those that would allow for interesting comparisons to be made. Residual variance (V_ϵ) was calculated as

$1 - (V_A + V_{Mg} + V_{Me} + COV_{A,Mg})$ (Willham, 1972).

We first simulated scenarios with the same total V_M, but varying in the proportion of maternal variation that was genetic. To do this we created three scenarios with no V_A and a total V_M of 0.5, with either all (scenario a), half (scenario b) or none (scenario c) of the maternal variation being genetic. To show the effect of the presence/absence of both direct genetic and maternal environmental effects, we simulated a scenario with only V_{Mg} (0.25; scenario d) and another with the same amount of V_{Mg} and additionally V_A (0.25) and V_{Me} (0.25; scenario e). To show the impact of $COV_{A,Mg}$, we simulated scenarios with maternal genetic and additive genetic variance with varying magnitude and direction of maternal and direct genetic covariance (no $COV_{A,Mg}$ in scenario f, positive $COV_{A,Mg}$ in scenarios g and h and negative $COV_{A,Mg}$ in scenarios i and j). For the sake of completeness, we also simulated two additional scenarios that we present in the supplements, with V_A but no V_{Mg}, and either with or without V_{Me} (scenarios k and l, see Figure S7).

Table 3: Simulated scenarios.

Scenario	V_A	V_{Mg}	V_{Me}	$COV_{A,Mg}$	$r_{A,Mg}$
a	0	0.5	0	0	0
b	0	0.25	0.25	0	0
c	0	0	0.5	0	0
d	0	0.25	0	0	0
e	0.25	0.25	0.25	0	0
f	0.25	0.25	0	0	0
g	0.25	0.25	0	0.075	0.3
h	0.25	0.25	0	0.15	0.6
i	0.25	0.25	0	-0.075	-0.3
j	0.25	0.25	0	-0.15	-0.6
k	0.25	0	0	0	0
l	0.25	0	0.25	0	0

As we were specifically interested in the bias in estimates from 'simple' maternal effects models, each dataset was analysed using an animal model estimating $\hat{V}_A$ and $\hat{V}_{Mc}$ as follows

$$z_i = \beta_0 + a_i + m_{C,j} + \epsilon_i \quad (3)$$

, where phenotype z of individual i is affected by its breeding value a_i, a maternal identity effect m_C of mother j and a residual ϵ_i, and β_0 represents the global mean. a, m_C and ϵ were all assumed to be normally distributed as follows:

$$\boldsymbol{a} \sim \mathcal{N}(0, \hat{V}_A \boldsymbol{A})$$
$$\boldsymbol{m_C} \sim \mathcal{N}(0, \hat{V}_{Mc} \boldsymbol{I})$$
$$\boldsymbol{\epsilon} \sim \mathcal{N}(0, \hat{V}_{\epsilon} \boldsymbol{I})$$

, where $\boldsymbol{A}$ is the relatedness matrix based on the pedigree and $\boldsymbol{I}$ is an identity matrix. All models were run using ASReml-R (version 4.1.0 Butler *et al.*, 2017). Using the results of these models, we estimated the bias in $\hat{V}_A$, $\hat{V}_M$ (total maternal variance) and $\hat{V}_{A_t}$. The bias was calculated for each combination of the 12 scenarios and 36 pedigree types. Bias was defined as $\frac{1}{n}\sum(\hat{\theta}_k - \theta)$ (where θ is the true value, $\hat{\theta}_k$ is the model estimate from kth simulation in a parameter set, and n is the number of simulations of that parameter set, i.e., 100). For V_M, θ was calculated as the sum of the simulated V_{Me} and V_{Mg}, and $\hat{\theta}_k$ was calculated using model estimates of $\hat{V}_{Mc}$. For V_{A_t}, θ was calculated as $V_A + \frac{3}{2}COV_{A,Mg} + \frac{1}{2}V_{Mg}$ and $\hat{\theta}_k$ was calculated using model estimates of $\hat{V}_A$.

We compared the bias for each scenario-pedigree combination to the mean proportion of non-sibling maternal links for that pedigree type. Although this metric specifically describes the links that contain unmodelled V_{Mg}, it is also informative for the amount of covariation shared due to $COV_{A,Mg}$ as the portion of informative links for both is highly correlated (Figure S5).

We also extracted the sampling covariance of $\hat{V}_A$ and $\hat{V}_{Mc}$ from the models. Sampling covariance gives information about how well the model can independently estimate the two variances (i.e. the identifiability). We looked to see whether the sampling covariance was predictive of the bias, or indicated the risk of bias - these results are presented in the Supplementary

materials (Section S6).

Modelling $\hat{V}_{Mg}$ as a solution

The results from the simulations above will inform us about the risk of bias in $\hat{V}_A$ when accounting for maternal effects by modelling $\hat{V}_{Mc}$. We also wanted to investigate whether we might be able to mitigate this bias. The clearest solution would be to run a full maternal effect animal model, and estimate $\hat{V}_{Mg}$ and $\hat{COV}_{A,Mg}$ regardless of the power we have to detect them. As discussed in the introduction, maternal genetic effects are often not modelled, likely because there is an assumption that estimating them would require a large and deep pedigree, as used in the simulations above, rather than the pedigrees typically available. Indeed, Meyer (1992) showed that the SE increased dramatically between models just estimating $\hat{V}_A$ and full maternal effect models estimating $\hat{V}_A$, $\hat{V}_{Me}$, $\hat{V}_{Mg}$ and $\hat{COV}_{A,Mg}$ (see Table 3 in Meyer 1992), although this appears at odds with results presented in Clément *et al.* 2001. In both cases, the simulations used what would be considered large pedigrees in the context of wild populations.

To assess the feasibility of running a full maternal effects model on smaller pedigrees, Author 1 ran an additional set of simulations. We wanted to create scenarios in which modelling maternal genetic effects would be challenging. To this end, we simulated two pedigree sizes. Based on the studies in Young & Postma (2023a), the median pedigree size used with an animal model was 420 individuals and the lower 10 and 25% quantiles were 105 and 175 individuals, respectively (Young & Postma, 2023b). We therefore simulated a small pedigree with 20 breeding females in each of two generations (the minimum needed to estimate maternal genetic effects), and a medium pedigree with 30 breeding females in each of four generations. Pedigrees were simulated with low fecundity (three offspring per female) and an intermediate mating system value, across the four immigration parameters, as these varied both in the proportion of non-sibling maternal links (Figure S4), and in pedigree quality (i.e the amount of missing parentage). This resulted in the small pedigree having 120 individuals with phe-

notypes (160-170 individuals in total), and the medium pedigree having 360 individuals with phenotypes (420-465 individuals in total). We then simulated 100 datasets for each of the 12 scenarios outlined in Table 3, for each of the 8 simulated pedigrees.

For each simulated data set, Author 1 ran four models; Model 1 included only additive genetic effects (allowing comparison to the results of Meyer, 1992):

$$z_i = \beta_0 + a_i + \epsilon_i \tag{4}$$

, with a and ϵ being normally distributed as outlined above (equation 3). Model 2 was the simple maternal effects model above (equation 3). Model 3 separated maternal genetic (m_G) and maternal environment effects (m_E), by additionally estimating maternal effects that were linked to the pedigree:

$$z_i = \beta_0 + a_i + m_{G,j} + m_{E,j} + \epsilon_i \tag{5}$$

, where

$$\boldsymbol{m_G} \sim \mathcal{N}(0, \hat{V}_{Mg}\boldsymbol{A})$$
$$\boldsymbol{m_E} \sim \mathcal{N}(0, \hat{V}_{Me}\boldsymbol{I})$$

. In this model $COV_{A,Mg}$ was assumed to be 0. Model 4 has the same structure as model 3, but additionally estimated $\hat{COV}_{A,Mg}$:

$$z_i = \beta_0 + a_i + m_{G,j} + m_{E,j} + \epsilon_i \tag{6}$$
$$[\boldsymbol{a}, \boldsymbol{m_G}] \sim N(\boldsymbol{0}, \begin{bmatrix} \hat{V}_A\boldsymbol{A} & \hat{COV}_{A,Mg}\boldsymbol{A} \\ \hat{COV}_{A,Mg}\boldsymbol{A} & \hat{V}_{Mg}\boldsymbol{A} \end{bmatrix})$$

.

To assess how well the different models performed, we calculated several metrics. We first calculated bias as outlined above. We also calculated precision as $1/\sqrt{\frac{1}{n}\sum(\hat{\theta}_k - \bar{\hat{\theta}})^2}$, where $\bar{\hat{\theta}}$ is the mean of the the model estimates from a parameter set. Because variance estimates are limited by 0, the standard deviation of the sampling distribution will decease as effect size nears zero, giving the appearance that precision decreases as effect sizes increase (Pick *et al.*, 2023). To account for this we also calculated relative precision as $\bar{\hat{\theta}}/\sqrt{\frac{1}{n}\sum(\hat{\theta}_k - \bar{\hat{\theta}})^2}$, which also represents the expected z-value. Finally, we calculated the Mean Absolute Error as $\frac{1}{n}\sum|\hat{\theta}_k - \theta|$. This is a measure of accuracy, combining both bias and precision, and represents the absolute deviation from true value. Note that this is a very similar measure to root mean squared error (RMSE), and our results are not affected by which measure of accuracy we use (the two measures have a correlation of $>$0.98 in our results; see Figures S15 and S16). Whilst we would expect that modelling $\hat{V}_{Mg}$ and $C\hat{O}V_{A,Mg}$ would address the issue of bias in $\hat{V}_A$, these models may increase the uncertainty, and so reduce the overall accuracy. Considering a measure of accuracy allows us to incorporate both when comparing the models. We therefore focus on this last metric in the results, and fully present all metrics in the Supplementary Materials.

Results

The extent of $\hat{V}_{Mg}$ across systems

From our survey of published estimates from wild animal populations, the overall mean proportion of phenotypic variation due maternal genetic effects ($\hat{m}^2_g$) was 0.143 (first year and older combined), and the mean for first year traits was 0.208 (Figure 2a, Table S1). Note that these averages do not consider sampling variation in the individual estimates (see Methods). These values are similar to those estimated across animal species in Moore *et al.* (2019b). The proportion of $\hat{V}_M$ due to $\hat{V}_{Mg}$ was generally high (mean=0.662), and again higher for

first year phenotypes (mean=0.720; Figure 2b). Two of the three very low values for this proportion come from estimates with low $\hat{m}^2$ (see Figure S3). In most cases, $\hat{V}_A$ was higher, and $\hat{V}_M$ lower, in the simple maternal effects models than when $\hat{V}_{Mg}$ was additionally estimated (Figure 2c and d, red bars). Interestingly, in the two cases with a notable decrease in heritability between simple and complex maternal effect models (grey bars in Figure 2c), a negative $\hat{COV}_{A,Mg}$ was also estimated in the complex model (see Table S1).

Non-sibling maternal links

Our simulations produced a large amount of variation in the proportion of non-sibling maternal links (defined in the methods and Table 1), that was similar to the range that might be encountered in common pedigrees of wild populations (Figure 3). Comparing pruned pedigrees for adults and juvenile phenotypes generated from the two wild pedigrees, we can also see that the pedigree structure can change dramatically between juvenile and adult phenotypes, with this difference being particularly pronounced in the avian system.

Bias in $\hat{V}_A$ and $\hat{V}_M$

Figure 4 shows the bias in $\hat{V}_A$ and $\hat{V}_M$ across the different scenarios, and Figure 5 the relationship between the two (full simulations results are shown in Figures S18-S29). When considering scenarios with no $COV_{A,Mg}$ (Figure 4A-D), we can see that as the proportion of non-sibling maternal links (see Table 1) increased, $\hat{V}_A$ from the simple maternal effects model became increasingly upwardly biased (Figure 4A and C), and estimates of maternal variance correspondingly downwardly biased (Figure 4B and D). There was a clear correspondence between the bias in $\hat{V}_A$ and the bias in $\hat{V}_M$, with the bias in $\hat{V}_M$ being approximately half the magnitude of the bias in $\hat{V}_A$ (solid line in Figure 5A). Although the proportion of non-sibling maternal links was clearly a strong predictor of the bias in $\hat{V}_A$ and $\hat{V}_M$, there was still additional unexplained variation in the bias caused by variation in pedigree structure (i.e., the scatter around predicted lines in Figure 4).

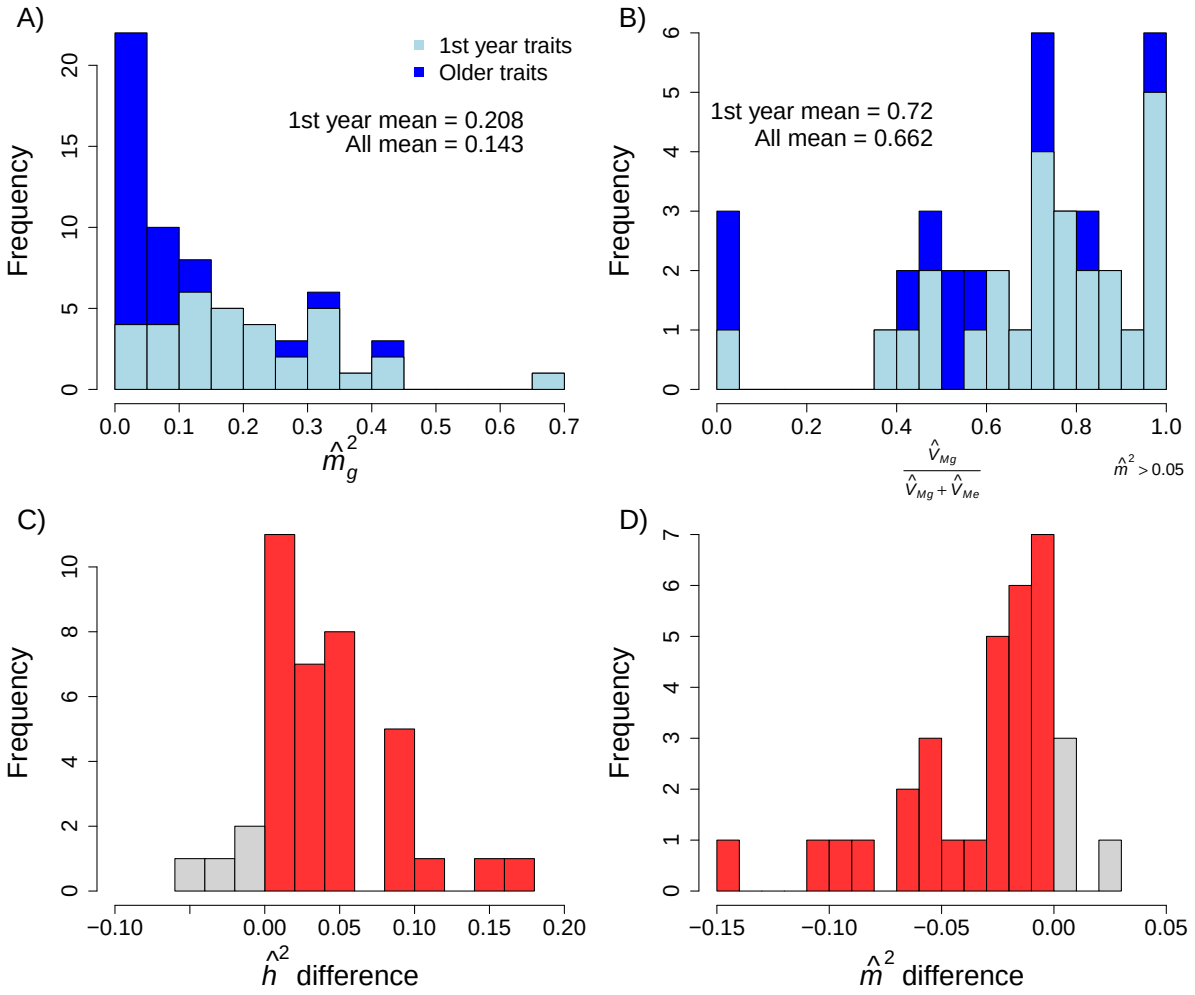


Figure 2: A) The proportion of total phenotypic variation due to maternal genetic effects from published estimates (i.e., $\frac{\hat{V}_{Mg}}{\hat{V}_P}$; n= 14 studies; see Table S1). B) The proportion of total maternal variation due to maternal genetic effects (i.e., $\frac{\hat{V}_{Mg}}{\hat{V}_{Mg}+\hat{V}_{Me}}$). This is subset for those estimates where the total proportion of $\hat{V}_M$ is above 0.05. C) and D) show the difference between $\hat{h}^2$ (C) and $\hat{m}^2$ (D) estimated in simple and complex maternal effect animal models. Red bars are those in which $\hat{h}^2$ was larger and $\hat{m}^2$ smaller in the simple model (estimating $\hat{V}_A$ and $\hat{V}_{Mc}$) than in the full model (estimating $\hat{V}_A$, $\hat{V}_{Mg}$ and $\hat{V}_{Me}$), which suggests overestimation of $\hat{h}^2$ and underestimation of $\hat{m}^2$ in the simple model.

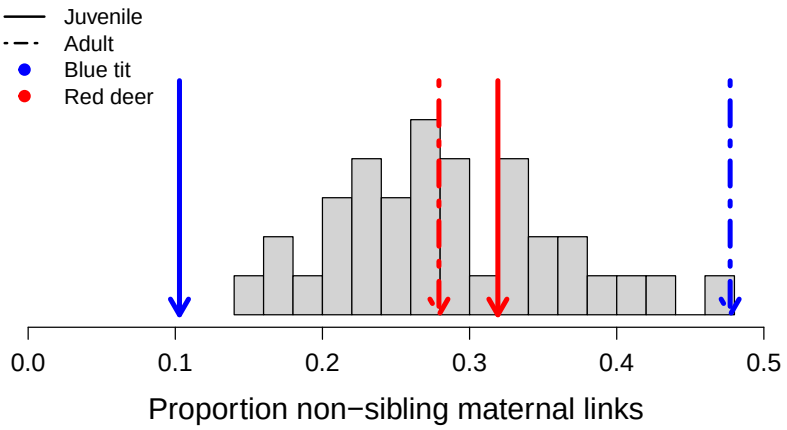


Figure 3: The distribution of non-sibling maternal links in simulated pedigrees (grey histogram) compared to observed pedigrees from two study systems red deer (red lines) and blue tits (blue lines). For each of the two real pedigrees, we generated a juvenile pedigree (solid lines), assuming that all individuals with a mother had a phenotype, and an adult pedigree (dashed lines), assuming only those that became parents (recruited) had a phenotype.

Scenarios a, b and c had the same amount of total maternal variance but vary from all to none of the variance being genetic. The comparison of these scenarios (Figure 4A and B) showed that the effect of proportion of non-sibling maternal links on the bias in $\hat{V}_A$ is dependent on the proportion of the total maternal variance that is genetic, with little or no bias when all maternal variance is environmental (blue diamonds), to a large bias in pedigrees with a high proportion of non-sibling maternal links when all maternal variance is genetic (black circles). Note that in these scenarios no V_A was simulated. It should also be noted that the small bias that can be seen in some pedigree structures in Scenario c, when the simulated maternal variance is environmental (blue diamonds), is due to estimated variances being upwardly biased when effect sizes are small (simulated V_A is 0 in this case), as variances are bound by 0 (see for example Pick *et al.*, 2023, see also Figure S20). Scenarios l and k show that, as expected, there was no bias in $\hat{V}_A$ in the absence of V_{Mg} (Figure S7). Comparison of scenarios d and e (Figure 4 C and D) showed the presence of both V_{Me} and V_A (yellow inverse triangles) decrease the bias in $\hat{V}_A$ caused by unmodelled V_{Mg}.

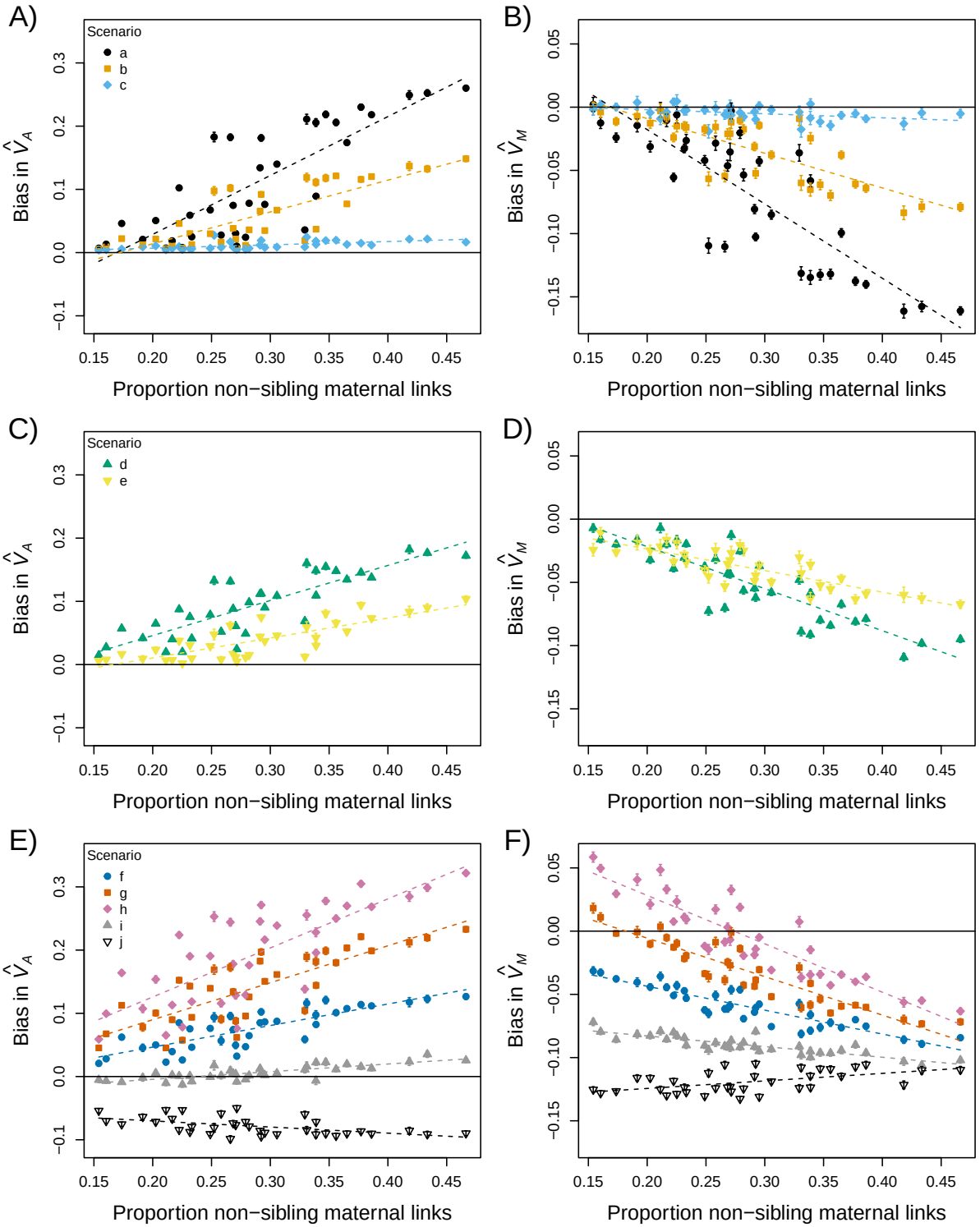


Figure 4: Bias in $\hat{V}_A$ (first column; A, C and E) and $\hat{V}_M$ (second column; B, D and F) in relation to the proportion of non-sibling maternal links across different simulated scenarios (indicated by the colours; see Table 3) and pedigree structures. The top row (A and B) compares scenarios with the same total V_M but different proportions of V_{Mg}. The second row (C and D) compares a scenario with just V_{Mg} with one that has the same V_{Mg} and additionally V_{Me} and V_A. The bottom row (E and F) shows the impact of different directions and magnitudes of $COV_{A,Mg}$. Dotted lines are predictions from a simple linear model, the purpose of which is just to help illustrate the pattern. Error bars show the standard error across simulations. Note that for some simulations the errors bars are too small to see.

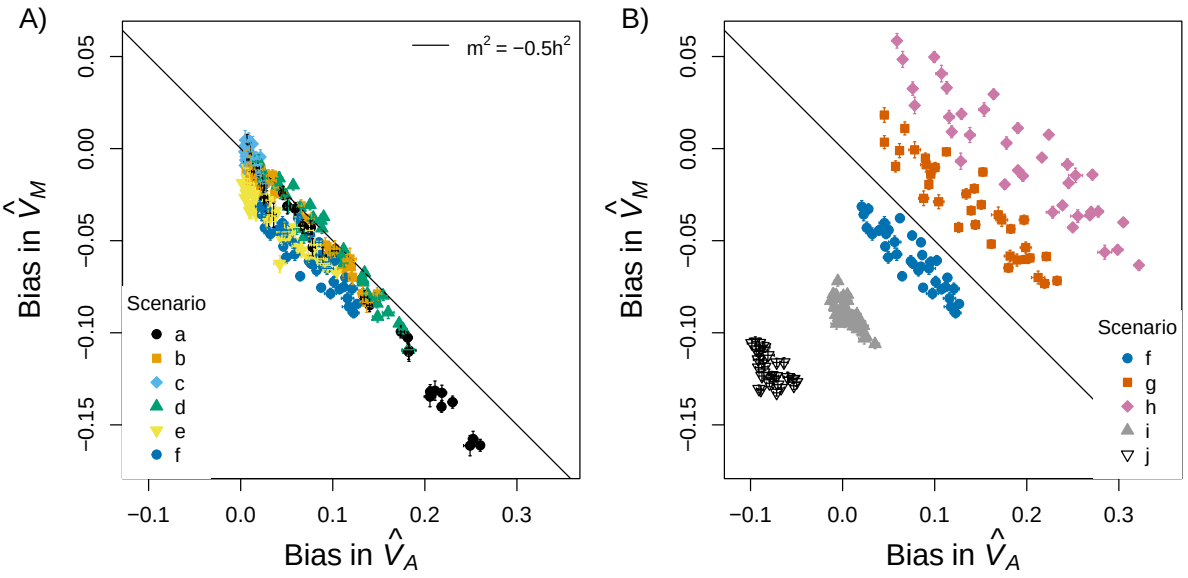


Figure 5: The relationship between the bias in $\hat{V}_A$ and $\hat{V}_M$ across different simulated scenarios (indicated by the colours and symbols; see Table 3) and pedigree structures. A) compares scenarios with no simulated $COV_{A,Mg}$, and B) shows the impact of different directions and magnitudes of simulated$COV_{A,Mg}$ on this relationship. Error bars show the standard error across simulations. The black line not based on theoretical predictions, rather is there for scale.

Effect of $COV_{A,Mg}$ on bias in $\hat{V}_A$ and $\hat{V}_M$

The comparison of scenario f (no $COV_{A,Mg}$) with scenarios g, h, i and j (Figure 4E and F) showed the impact of $COV_{A,Mg}$ on the bias in $\hat{V}_A$ and $\hat{V}_M$. The effect of non-sibling maternal links on the bias was clearly dependent on the presence and direction of the covariance. When the covariance was positive the bias in $\hat{V}_A$ was increased with increasing non-sibling maternal links (pink diamonds and red squares), whilst a negative covariance (grey triangles and open inverse triangles) reduced the effect of non-sibling maternal links on $\hat{V}_A$, and even changed the direction of the bias, leading to an underestimation of $\hat{V}_A$ when the covariance was moderately negative (r=0.6), which is in line with the results in Figure 2c.

In contrast to the scenarios without $COV_{A,Mg}$, the relative amount of bias in $\hat{V}_A$ and $\hat{V}_M$ was strongly affected by the presence of a genetic covariance (Figure 5B). A negative covariance resulted in relatively more bias in $\hat{V}_M$ (grey and open points in Figure 5B), whereas a positive covariance increased the bias in $\hat{V}_A$ relative to $\hat{V}_M$ (red and pink points in Figure 5B).

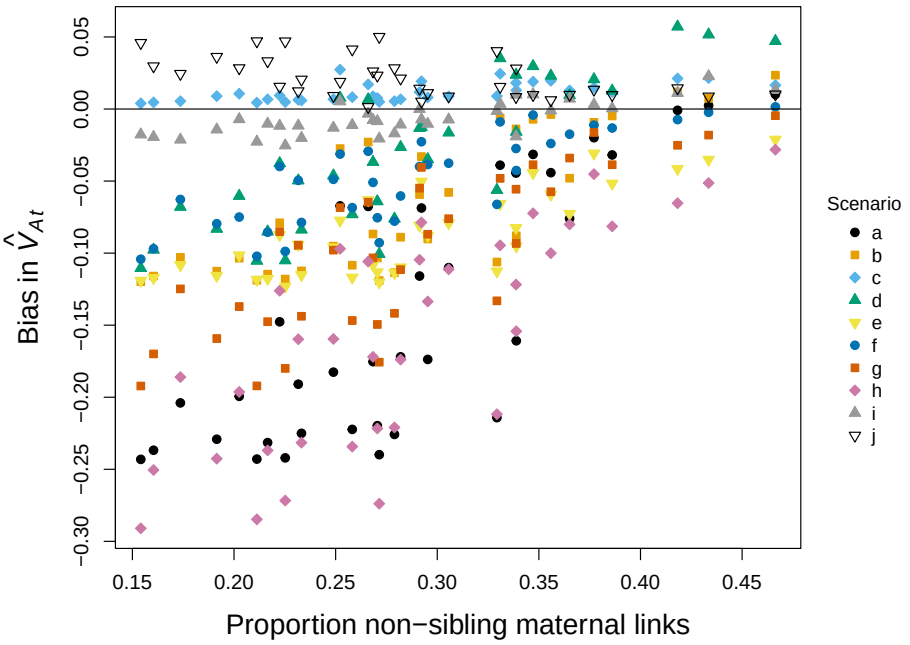


Figure 6: Bias in $\hat{V}_{A_t}$, from simulations with different underlying parameters (indicated by the colours; see Table 3) and different pedigree structures.

Estimating evolutionary potential

As we might expect, evolutionary potential (measured as $\hat{V}_{A_t}$) was generally (but not always) underestimated across scenarios when using simple maternal effects model to estimate $\hat{V}_A$ (Figure 6). However, the underestimation was lower when the proportion of non-sibling links in the pedigree was higher. When there was negative $COV_{A,Mg}$, the bias was minimal (at least under the parameter values simulated here), and in the most extreme scenario even positive (open inverse triangles in Figure 6).

Modelling $\hat{V}_{Mg}$ in small pedigrees

The results from simulations of small and medium pedigrees were qualitatively very similar across the different immigration parameters and pedigree sizes (Table S2) and so we focus on the results from the small pedigrees with unbiased immigration in Figure 7. As expected, the medium size pedigrees were generally more precise, and also less biased in several scenarios (Figures S11- S14).

The bias in $\hat{V}_A$, $\hat{V}_M$ and $\hat{V}_{A_t}$ across scenarios was centred on zero only for model 3 (Figures S11- S13). $\hat{V}_A$, $\hat{V}_M$ and $\hat{V}_{A_t}$ in model 4 were biased under many scenarios, but the bias decreased with increasing pedigree size (Figures S11- S13), whereas it was largely unchanged in the other models. Model 4 also showed biases in $\hat{V}_{Mg}$ and $\hat{COV}_{A,Mg}$, especially when there was no underlying V_A (for $\hat{COV}_{A,Mg}$) or V_{Mg} (both estimates), and again these biases decreased with increasing pedigree size (see Figure S14, and discussed further in the Supplementary material). Most notably, there was a clear bias towards negative covariances when V_A or V_{Mg} were absent. Interestingly, precision was broadly similar across models 1-4 for $\hat{V}_A$, $\hat{V}_M$ and $\hat{V}_{A_t}$ (S11- S13).

Consequently, models that estimate maternal genetic effects were generally more accurate (measured as mean absolute error, a combination of bias and precision) in terms of $\hat{V}_A$ estimation than simple maternal effects models (Figure 7, top row). They also displayed similar levels of accuracy for the estimation of $\hat{V}_M$ and $\hat{V}_{A_t}$ (Figure 7, middle and bottom rows). This indicates that there appears to be no clear cost to estimating maternal genetic effects, at least under the conditions simulated here.

Model 4 (estimating $\hat{COV}_{A,Mg}$) had some convergence issues, with around 9% of models not converging overall (compared to 0% for the other 3 models). This was especially problematic in simulated scenarios where there was no simulated V_A (scenarios a-d), where approx 15-25% of the models failed to converge (Figure S17). In these scenarios without V_A (scenarios a-d), model 3 also outperformed model 4 (middle column in Figure 7), especially for $\hat{V}_A$ estimation. In scenarios where both V_A and V_{Mg} were simulated, it was either better or no different in terms of accuracy to estimate $\hat{COV}_{A,Mg}$ (first column in Figure 7). In combination with the biases mentioned above, care should be taken when interpreting estimates taken from model 4, especially with small pedigrees. We discuss this further in the supplementary material.

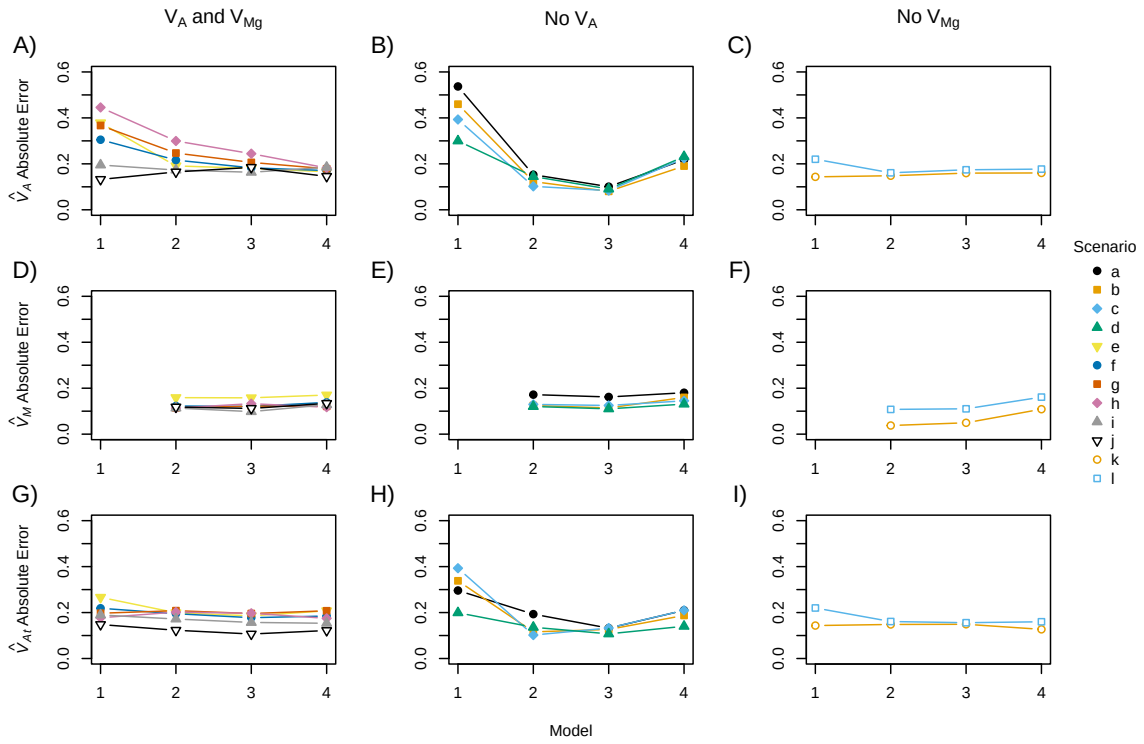


Figure 7: Accuracy (measured as mean absolute error) in estimation of $\hat{V}_A$ (A, B and C) $\hat{V}_M$ (D, E and F) and $\hat{V}_{A_t}$ (G, H and I), estimated using 4 different models: 1) estimated $\hat{V}_A$ only, 2) estimated $\hat{V}_A$ and $\hat{V}_{Mc}$, 3) estimated $\hat{V}_A$, $\hat{V}_{Mg}$ and $\hat{V}_{Me}$ and 4) estimated $\hat{V}_A$, $\hat{V}_{Mg}$, $\hat{V}_{Me}$ and $\hat{COV}_{A,Mg}$. Data from these plots is from simulations of small pedigrees with unbiased immigration, over 12 scenarios (shown with different symbols and colours), which are separate out across the three columns, to show scenarios with both V_A and V_{Mg}, no V_A and no V_{Mg}, respectively.

Discussion

Here we show that, based on available estimates, maternal variation is likely to have a considerable genetic component. Simple models of maternal effects (which dominate the literature) are biased in the presence of maternal genetic effects: the direct additive genetic variation (V_A) is likely commonly overestimated, and total maternal variation (V_M) underestimated. This occurs because the modelled maternal identity effects only account for the similarity between individual that share the same mother, but there are other individuals in the pedigree that additionally share maternal genetic variance (V_{Mg}). These biases are dependent on the underlying parameter values and the pedigree structure, and in particular the proportion of non-sibling maternal links; pedigrees with a high proportion of these links show high bias. The presence of direct-material genetic covariance ($COV_{A,Mg}$) also affects the bias, with the positive covariance increasing the bias and a negative covariance decreasing, or even reversing its direction. This bias in $\hat{V}_A$ additionally affects the estimation of $\hat{V}_{A_t}$. In simple maternal effects models, this is based solely on $\hat{V}_A$, and so will be systematically underestimated in the presence of V_{Mg}. The upward bias in $\hat{V}_A$ in simple maternal effects models therefore actually acts to reduce the bias in $\hat{V}_{A_t}$. However, the bias is not completely removed and so $\hat{V}_{A_t}$ is still typically underestimated, although this is dependent on the underlying levels of V_{Mg} and $COV_{A,Mg}$. Consequently, without fully modelling sources of direct and indirect genetic variation, we are limited in our understanding of the full evolutionary potential of a trait.

Maternal variation is likely to have a considerable genetic component (>50%, Figure 2b). We note that all but one species for which we found data were mammals. However, we believe that these results are likely to generalise. First, maternal effects are likely to occur to a similar extent in other taxa; Moore *et al.* (2019b) found no difference in the size of maternal effects between egg-laying and live-bearing species, vertebrate and invertebrate species, amniote or anamniote species, and species providing post-natal care or not. Second, this matches previous work looking at the proportion of consistent individual variation that is genetic (mean of 0.52 for behavioural traits: Dochtermann *et al.*, 2015). Given realistic pedigree structures, this level

of V_{Mg} will bias the estimation of $\hat{V}_A$ and of $\hat{V}_M$. Consequently, we expect that the average $\hat{h}^2$ and $\hat{m}^2$ presented in Moore *et al.* (2019b) are systematically over- and under-estimated, respectively. How meaningful is the bias likely to be, and what can we do about it?

Under intermediate levels of non-sibling maternal links, and realistic parameter values (such as those considered here), we may expect an upward bias in $\hat{h}^2$ in juvenile traits of 0.05-0.1 (note that this will be heavily dependent on underlying parameter values). Correspondingly, we found that $\hat{h}^2$ estimates were commonly larger in simple compared with complex maternal effect models by up to around 0.15 (Figure 2c). This represents a considerable proportional increase, relative to the average size of $\hat{h}^2$ found across studies (0.2-0.3 Postma, 2014; Moore *et al.*, 2019b). This level of bias in $\hat{V}_A$ would therefore result in quite substantially over-estimation of the predicted selection response, when using models such as the breeder's equation. Indeed, misestimation of V_A is commonly suggested as a reason for why our predictions of evolutionary change commonly do not match our observations ('the paradox of stasis' Merilä *et al.*, 2001; Pujol *et al.*, 2018).

Although these models will likely overestimate $\hat{V}_A$, $\hat{V}_{A_t}$ was typically underestimated under our simulated scenarios (Figure 6). If V_{Mg} is common, in juvenile traits at least, we will be commonly *underestimating* the evolutionary potential of these traits, as we are not explicitly considering indirect genetic effects. The use of V_{A_t} as a measure of evolutionary potential assumes, however, that selection is only directly acting on the juvenile trait, and not on maternal performance. Selection response in the juvenile trait is dependent not only on direct selection on that trait, but also on selection acting on the maternal traits and thus maternal performance (Cheverud, 1984; Kirkpatrick & Lande, 1989). Theoretically, we would predict that selection acts in the opposite direction on maternal performance than on the offspring trait, as maternal care is often predicted to be costly (Cheverud, 1984; Hadfield, 2012). Only two studies to date have directly estimated selection on maternal performance (Thomson *et al.*, 2017; Gauzere *et al.*, 2022), with opposing results. Dependent on the underlying variances and covariances, selection on maternal performance could produce a strong enough force to

constrain the response to selection on the juvenile trait. In such situations, it is important to know where the genetic variation is coming from to correctly predict selection response. In situations where $\hat{V}_{A_t}$ is well estimated by simple maternal effects models (e.g. when there is a high proportion of non-sibling maternal links), $\hat{V}_A$ is overestimated and $\hat{V}_M$ is underestimated when there is V_{Mg}. Without knowledge of these biases, it may be assumed that selection on maternal performance is not important because there are a negligible maternal effects compared to the amount of direct genetic variation. However, although $\hat{V}_{A_t}$ is well estimated in these conditions, if there were selection on maternal performance, the response to selection would be much different than predicted. In other words, situations in which we might wrongly dismiss maternal effects are also the scenarios in which correctly characterising the source of the genetic variation is particularly important. It is therefore important to try and separate out the different sources of genetic variation, where possible.

Interestingly, the biases in $\hat{V}_A$ and $\hat{V}_M$ may change dramatically for different traits within the same study system (Figure 3). This is because the proportion of non-sibling maternal links can change substantially between pedigree for juvenile and adult traits. For example, in species with high fecundity and low recruitment (for example, passerine birds), much of the pedigree information for juvenile traits will come from comparisons between siblings, for whom the maternal variance is modelled. For adult traits in these systems, the pedigree information largely comes from parent-offspring comparisons, half of which will be maternal (and perhaps more, in systems where it is more likely that the mother is known). Maternal effects are generally found to decrease in adult phenotypes (Wilson & Reale, 2006; Pick *et al.*, 2016a; Moore *et al.*, 2019b; Gauzere *et al.*, 2020b). Although this may be expected, as the intensity of interactions between mothers and offspring reduces over time, this decrease may also be accentuated by the change in pedigree structure, as the overestimation of V_A and underestimation of V_M might be stronger at this stage.

When considering the impact of these biases it is also worth considering that the error associated with $\hat{V}_A$ is typically large. The mean standard error of $\hat{h}^2$ from animal models based on

syntheses is around 0.1 (0.099 from Postma 2014, 0.097 from Young & Postma 2023a calculated using data from Young & Postma 2023b, and 0.115 from Moore *et al.* 2019b calculated using data from Moore *et al.* 2019a), and so the difference in $\hat{h}^2$ between simple and complex maternal effects models is likely to fall within the confidence intervals of $\hat{h}^2$. Consequently, for any single study, this bias might not alter the inference too much (assuming the estimate is being interpreted in the context of the confidence intervals). Syntheses (such as Postma, 2014; Moore *et al.*, 2019b) will, however, systematically overestimate V_A.

To address these biases in the estimation of $\hat{V}_A$ and $\hat{V}_M$, we would ideally run full maternal genetic effect models. This would provide us with unbiased estimates of $\hat{V}_A$, $\hat{V}_M$ and $\hat{V}_{A_t}$, as well as $\hat{V}_{Mg}$ and $\hat{COV}_{A,Mg}$. As discussed above, it is particularly useful to have these latter parameters, as selection may act in the opposite direction on maternal performance than on the offspring trait, if maternal care is costly (Cheverud, 1984; Hadfield, 2012). To fully predict selection response, these parameters must therefore be separately estimated. Full maternal effect animal models are seldom run, however. This may be due to a perception that these models require restrictively large sample sizes to run, and so this may not be realistic in many cases. However, our simulations show that for even small pedigrees, running full maternal effects animal models (or at least those additionally estimating $\hat{V}_{Mg}$) leads to the highest accuracy in $\hat{V}_A$, and no loss of accuracy for $\hat{V}_{A_t}$ and $\hat{V}_M$, with the additional advantage of separating genetic and environmental sources of maternal variation. This matches the results shown in Clément *et al.* 2001, although that study was based on much larger pedigrees. Depending on the underlying parameters, modelling $\hat{COV}_{A,Mg}$ may prove challenging. In simulations without V_A, 15-25% of models estimating $\hat{COV}_{A,Mg}$ failed to converge and were more biased than models that did not estimate $\hat{COV}_{A,Mg}$ especially at small pedigree sizes, showing evidence of identifiability issues when V_A or V_{Mg} were not present. However, our simulations suggest that estimating $\hat{V}_{Mg}$ and $\hat{V}_{Me}$ in addition to $\hat{V}_A$ would provide an increase in accuracy and understanding of the underlying processes (although we note that they would require increased computational power). We therefore recommend that these models are more

frequently used and applied to an extended range of systems.

If running full maternal genetic effect models is not possible, then enough information about the pedigree needs to be presented to allow the risk of bias to be assessed. Pedigree metrics (for example, those generated by the R package pedantics (now maintained as pedtricks); Morrissey & Wilson, 2010; Martin *et al.*, 2024) are often reported alongside animal models. However, the explicit utility of these metrics has not been explored (i.e., whether and how they relate to precision and/or bias in estimates). Here we show that the proportion of non-sibling maternal links provides a good prediction for the potential for bias in $\hat{V}_A$ and $\hat{V}_M$ from unmodelled $\hat{V}_{Mg}$, although we note that this metric does not completely predict the bias (see Figure 4). There are likely additional metrics (e.g., relating to the relative amount of maternal siblings) that would explain an additional amount of variation. Importantly, this metric gives no information about the actual bias, as this is dependent on the underlying parameters, which are unknown. Given that it explains most of the variation in the potential for bias, we recommend reporting this metric alongside simple maternal effects models, so that the potential for bias in $\hat{V}_A$ and $\hat{V}_M$ can be assessed. To fully explore the potential for bias in a given pedigree, we recommend that a simulation approach is taken, similar to that taken here, but focussing on that single pedigree structure.

The maternal effect model presented here represents the variance partitioning approach to estimating maternal effects. This method can be extended by additionally including maternal phenotypes in the model (trait-based approach) to show the extent to which they explain the maternal variance (referred to as the Hybrid trait-based/variance component approach Hadfield, 2012; McAdam *et al.*, 2014). This approach has successfully been used in several systems (Hadfield *et al.*, 2013; Noble *et al.*, 2014; Pick *et al.*, 2016a; Gauzere *et al.*, 2021). The biases we demonstrate here have interesting consequences in this context. If we take the example of system where a single maternal phenotype explains all the maternal variation. If the maternal phenotype in question has a considerable genetic component (which would generate V_{Mg}), a simple maternal effect model will underestimate $\hat{V}_M$, and over-estimate

$\hat{V}_A$. Adding the maternal phenotype into the model will decrease the estimates of maternal variation to 0, but it will also reduce $\hat{V}_A$, as the maternal genetic variance that was upwardly biasing it is now explained. We can see this exact effect in the example of egg size in Japanese quail (*Coturnix japonica*). Maternal egg size has considerable genetic component (Pick *et al.*, 2016b, 2019b), and so there would be considerable V_{Mg} in any juvenile traits that it affects. Pick *et al.* (2016a) found that including maternal egg size in a simple maternal affects model of hatching size reduced both $\hat{V}_A$ and $\hat{V}_M$ to effectively zero (see Figure 2 in Pick *et al.*, 2016a). This indicates that the $\hat{h}^2$ estimated in the simple maternal effects model (0.268) was an artefact of unmodelled V_{Mg}, and suggests that, in cases where the maternal phenotypes driving the maternal effects are known, the hybrid approach may provide a useful way to reduce bias in $\hat{V}_A$. We note that the presence of a genetic covariance between maternal and offspring traits will complicate this, and may still lead to an over or under estimation of $\hat{V}_A$ in a hybrid model, depending on the direction of the covariance. Multivariate models may therefore be more appropriate where there is evidence of $COV_{A,Mg}$.

Here we focus specifically on maternal effects, which are an important source of variation in species with uniparental care. The large focus on maternal effects in the quantitative genetics literature is likely due to the majority of species used in animal breeding having uniparental, maternal care (i.e., cows, sheep, pigs, chicken). Similarly, all published estimates of $\hat{V}_{Mg}$ in the wild come from mammals exhibiting *maternal* care. However, many commonly studied systems in the wild (e.g passerine birds) have biparental care. In these systems, there is likely both maternal and paternal effects. What impact would these paternal effects have?

In avian systems, we commonly model nest effects to capture the joint parental effects in a particular reproductive attempt. Like modelling maternal identity, these nest effects will likely not fully capture the maternal and parental genetic variation. Under certain assumptions, the impact of paternal genetic effects would therefore be expected to be the same (i.e., unmodelled paternal genetic variation would upwardly bias $\hat{V}_A$), but the bias would be linked to the proportion of non-sibling *paternal* links. There are several complications, however. In

populations with extra-pair paternity, individual's will not always be raised by their genetic father, which dilutes the confounding between paternal genetic and additive genetic effects. We would therefore expect less confounding between paternal and direct genetic effects in genetic than social pedigrees, and for the bias to depend on both pedigree structure and the extend of extra-pair paternity. This situation would be further complicated by the presence of any genetic covariance between maternal and paternal effects, and any relatedness between parents. These issues would make for an interesting extension to this study, and Varona *et al.* (2015) have explored similar issues in the context of paternal imprinting.

Cross fostering is also a method used to help disentangle genetic and (post-natal) common environment effects in avian systems. Whole brood cross fostering (swapping whole litter/broods between nests) should get rid of the biases caused by both parental environmental and genetic effects, as the genetic parents no longer rear the offspring; for example, two maternally related cousins wouldn't be raised by related mothers, and so wouldn't also share V_{Mg}. This method, however, has limited power to separate parental and genetic effects generally, especially in the absence of a genetic pedigree. It also assumes that parental effects only occur after cross fostering; any parental effects occurring before crossing takes place (e.g., pre-natal maternal effects) would still bias $\hat{V}_A$. In partial cross fostering (where some chicks remain in the nest of origin and some are moved), on the other hand, the chicks that are not crossed still receive care from their genetic mother. A back of the envelope calculation would suggest that 25% of the bias in $\hat{V}_A$ would remain if 50% of the offspring were crossed (the bias would remain for any two maternal relatives that were raised by their mother, and each would have a 50% chance of being crossed, meaning 25% of the maternal relations would both be raised by their genetic mother). This would clearly need further investigation, but it suggests that while partial cross fostering presents a more powerful approach for accounting for common environment effects, it does not fully account for parental genetic effects. This would explain some results shown in Kruuk & Hadfield (2007, Table 3). In collared flycatchers, both $\hat{V}_A$ and $\hat{V}_P$ from animal models on juvenile body mass and condition were substantially reduced

in cross-fostered chicks. The decrease in phenotypic variance was attributed to the potential presence of $COV_{A,Mg}$, but the reduction in $\hat{V}_A$ was unexplained. This reduction may therefore be the result of the confounding of direct and parental genetic effects being broken up in the cross fostered chicks.

Conclusions

It is well established that maternal effects (and more generally common environment effects) can affect our estimation of V_A, which is a key target for estimation in quantitative genetic studies. Our study shows that the common methods for accounting for maternal effects do not fully do so, meaning that under commonly seen levels of maternal and direct genetic variance and pedigree structures, our estimation of genetic variation is biased. The inference about the evolutionary potential that we can make from these simple maternal effects models is therefore limited. Our simulations also show that models explicitly estimating maternal genetic effects are no less accurate through greater imprecision, even when pedigrees are small (although we note that inference about these parameters at small sample sizes is likely generally limited by high levels of uncertainty). We therefore recommend that maternal genetic effects are estimated if there is any evidence for the presence of maternal effects. Models that additionally estimate $\hat{COV}_{A,Mg}$ can be problematic when there is no underlying V_A or V_{Mg}, especially when pedigrees are small, and so we suggest to first run a model estimating $\hat{V}_A$, $\hat{V}_{Mg}$ and $\hat{V}_{Me}$, and then further estimating $\hat{COV}_{A,Mg}$ if there is evidence of both V_A and V_{Mg}. Care should be taken with the interpretation of the parameter estimates from these models (especially with negative $\hat{COV}_{A,Mg}$) when pedigrees are small, due to identifiability issues. We also recommend not dropping $\hat{V}_{Mg}$ from models if there is no statistical support for it (i.e., model simplification); the lack of statistical support does not indicate the lack of V_{Mg} (it more likely indicates the lack of power to detect it), and we find no effect of estimating $\hat{V}_{Mg}$ on accuracy. $\hat{V}_A$ will therefore be less biased and no less accurate when $\hat{V}_{Mg}$ is estimated. With small datasets we will always struggle to estimate genetic variances, but this appears

to be made no worse by using more complex models. As studies of maternal genetic effects are rare and taxonomically limited, more detailed modelling of maternal and paternal genetic effects in the wild would give greater insight into their evolutionary importance.

Acknowledgements

We thank Josephine Pemberton, Lizy Mittell and Kasha Strickland for constructive feedback. JLP and LEBK were funded by LEBK's Royal Society Professorship (RSRP-R1-211017) and ERC grant (101020503 — EVOWILD — ERC-2020-ADG).

Data and Code Availability

https://github.com/joelpick/maternal_effects.

Author Contributions

Joel Pick: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Visualization, Writing – original draft, Writing – review & editing

Craig Walling: Supervision, Writing – review & editing

Loeske Kruuk: Conceptualization, Funding acquisition, Supervision, Writing – original draft

Conflict of Interest statement

We declare no conflicts of interest

References

Adams, M.J., Robinson, M.R., Mannarelli, M.e. & Hatchwell, B.J. 2015. Social genetic and social environment effects on parental and helper care in a cooperatively breeding bird.

Proc. R. Soc. B: Biol. Sci. **282**:20150689.

Bell, A.M., Trapp, R. & Keagy, J. 2018. Parenting behaviour is highly heritable in male stickleback. *R. Soc. Open Sci.* **5**.

Beraldi, D., McRae, A.F., Gratten, J., Slate, J., Visscher, P.M. & Pemberton, J.M. 2007. Mapping Quantitative Trait Loci Underlying Fitness-Related Traits in a Free-Living Sheep Population. *Evolution*. **61**:1403–1416.

Bernardo, J. 1996. Maternal effects in animal ecology. *Am. Zool.* **105**:83–105.

Blomquist, G.E. & Williams, L.E. 2013. Quantitative genetics of costly neonatal sexual size dimorphism in squirrel monkeys (Saimiri boliviensis). *J. Evol. Biol.* **26**:756–765.

Bonnet, T., Morrissey, M.B., de Villemereuil, P., Alberts, S.C., Arcese, P., Bailey, L.D. *et al.* 2022. Genetic variance in fitness indicates rapid contemporary adaptive evolution in wild animals. *Science*. **376**:1012–1016.

Butler, D., Cullis, B., Gilmour, A., Gogel, B. & Thompson, R. 2017. ASReml-R Reference Manual.

Bérénos, C., Ellis, P.A., Pilkington, J.G. & Pemberton, J.M. 2014. Estimating quantitative genetic parameters in wild populations: a comparison of pedigree and genomic approaches. *Mol. Ecol.* **23**:3434–3451.

Cheverud, J. & Moore, A.J. 1994. Quantitative Genetics and the Role of the Environment Provided by Relatives in Behavioral Evolution. In: *Quantitative genetic studies of behavioral evolution* (C.R.B. Boake, ed.), University of Chicago Press.

Cheverud, J.M. 1984. Evolution by kin selection: a quantitative genetic model illustrated by maternal performance in mice. *Evolution*. **38**:766–777, iSBN: 0014-3820.

Class, B. & Brommer, J.E. 2020. Can dominance genetic variance be ignored in evolutionary quantitative genetic analyses of wild populations? *Evolution*. **74**:1540–1550.

Clément, V., Bibé, B., Verrier, E., Elsen, J.M., Manfredi, E., Bouix, J. *et al.* 2001. Simulation analysis to test the influence of model adequacy and data structure on the estimation of genetic parameters for traits with direct and maternal effects. *Genet. Sel. Evol*. **33**:369–95.

Dickerson, G. 1947. Composition of hog carcasses as influenced by heritable differences in rate and economy of gain. *Iowa Agric. Exp. Stn. Res. Bull*. **354**:492–524.

Dochtermann, N.A., Schwab, T. & Sih, A. 2015. The contribution of additive genetic variation to personality variation: heritability of personality. *Proc. R. Soc. B: Biol. Sci*. **282**:20142201.

Dor, R. & Lotem, A. 2010. Parental effort and response to nestling begging in the house sparrow: Repeatability, heritability and parent-offspring co-evolution. *J. Evol. Biol*. **23**:1605–1612.

Falconer, D. 1981. *Introduction to Quantitative Genetics*. 2nd edn., Longman Group Ltd., London.

Freeman-Gallant, C.R. & Rothstein, M.D. 1999. Apparent heritability of parental care in Savannah Sparrows. *Auk* **116**:1132–1136.

Galloway, L.F., Etterson, J.R. & McGlothlin, J.W. 2009. Contribution of direct and maternal genetic effects to life-history evolution. *New Phytol*. **183**:826–838.

Gauzere, J., Pemberton, J.M., Kruuk, L.E.B., Morris, A., Morris, S. & Walling, C.A. 2022. Maternal effects do not resolve the paradox of stasis in birth weight in a wild red deer populaton. *Evolution*. **76**:2605–2617.

Gauzere, J., Pemberton, J.M., Morris, S., Morris, A., Kruuk, L.E. & Walling, C.A. 2020a. Data from: The genetic architecture of maternal effects across ontogeny in the red deer.

Gauzere, J., Pemberton, J.M., Morris, S., Morris, A., Kruuk, L.E. & Walling, C.A. 2020b. The genetic architecture of maternal effects across ontogeny in the red deer. *Evolution*. **74**:1378–1391.

Gauzere, J., Walling, C.A., Pick, J.L., Watt, K., Jack, P., Morris, A. *et al.* 2021. The role of maternally transferred antibodies in maternal performance in red deer. *Ecol. Lett*. **24**:2065–2076.

Hadfield, J.D. 2012. The quantitative genetic theory of parental effects. In: *The Evolution of Parental Care* (N.J. Royle, P.T. Smiseth & M. Kölliker, eds.), pp. 267–284, Oxford University Press.

Hadfield, J.D., Heap, E.A., Bayer, F., Mittell, E.A. & Crouch, N.M.A. 2013. Disentangling genetic and prenatal sources of familial resemblance across ontogeny in a wild passerine. *Evolution*. **67**:2701–13.

Henderson, C.R. 1988. Theoretical basis and computational methods for a number of different animal models. *J. Dairy Sci*. **71**:1–16.

Ibáñez, B., Cervantes, I., Gutiérrez, J.P., Goyache, F. & Moreno, E. 2014. Estimates of direct and indirect effects for early juvenile survival in captive populations maintained for conservation purposes: the case of Cuvier's gazelle. *Ecol. Evol*. **4**:4117–4129.

Kirkpatrick, M. & Lande, R. 1989. The evolution of maternal characters. *Evolution*. **43**:485–503.

Kruuk, L.E.B. 2004. Estimating genetic parameters in natural populations using the 'animal model'. *Phil. Trans. R. Soc. Lond. Ser. B: Biol. Sci*. **359**:873–890.

Kruuk, L.E.B. & Hadfield, J.D. 2007. How to separate genetic and environmental causes of similarity between relatives. *J. Evol. Biol*. **20**:1890–1903.

Lande, R. 1976. Natural Selection and Random Genetic Drift in Phenotypic Evolution. *Evolution*. **30**:314–334.

Lande, R. & Kirkpatrick, M. 1990. Selection response in traits with maternal inheritance. *Genet. Res.* **55**:189–197.

Lush, J.L. 1937. *Animal Breeding Plans*. Iowa State College Press, Ames, Iowa.

Lynch, M. 1987. Evolution of intrafamilial interactions. *Proc. Natl. Acad. Sci.* **84**:8507–8511, publisher: Proceedings of the National Academy of Sciences.

Lynch, M. & Walsh, B. 1998. *Genetics and Analysis of Quantitative Traits*. Sinauer Associates, Inc., Sunderland, MA.

Maccoll, A.D.C. & Hatchwell, B.J. 2003. Heritability of Parental Effort in a Passerine Bird. *Evolution*. **57**:2191–2195.

Martin, J., Wolak, M., Johnston, S. & Morrissey, M.B. 2024. pedtricks: Visualize, Summarize and Simulate Data from Pedigrees.

McAdam, A.G., Boutin, S., Réale, D. & Berteaux, D. 2002. Maternal effects and the potential for evolution in a natural population of animals. *Evolution*. **56**:846–851.

McAdam, A.G., Garant, D. & Wilson, A.J. 2014. The effects of others' genes: Maternal and other indirect genetic effects. In: *Quantitative Genetics in the Wild* (A. Charmantier, D. Garant & L.E.B. Kruuk, eds.), pp. 84–103, Oxford University Press, Oxford.

McFarlane, S.E., Gorrell, J.C., Coltman, D.W., Humphries, M.M., Boutin, S. & Mcadam, A.G. 2014. Very low levels of direct additive genetic variance in fitness and fitness components in a red squirrel population. *Ecol. Evol.* **4**:1729–1738.

McFarlane, S.E., Gorrell, J.C., Coltman, D.W., Humphries, M.M., Boutin, S. & Mcadam, A.G. 2015. The nature of nurture in a wild mammal's fitness. *Proc. R. Soc. Lond. B* **282**:1–7.

McGlothlin, J.W., Moore, A.J., Wolf, J.B. & Brodie III, E.D. 2010. INTERACTING PHENOTYPES AND THE EVOLUTIONARY PROCESS. III. SOCIAL EVOLUTION. *Evolution* **64**:2558–2574.

Merilä, J., Sheldon, B.C. & Kruuk, L.E.B. 2001. Explaining statis: Microevolutionary studies in natural populations. *Genetica* **112-113**:199–222.

Meyer, K. 1992. Bias and sampling covariances of estimates of variance components due to maternal effects. *Genet. Sel. Evol.* **24**:487.

Moore, A.J., III Brodie, E.D. & Wolf, J.B. 1997. INTERACTING PHENOTYPES AND THE EVOLUTIONARY PROCESS: I. DIRECT AND INDIRECT GENETIC EFFECTS OF SOCIAL INTERACTIONS. *Evolution*. **51**:1352–1362.

Moore, M.P., Whiteman, H.H. & Martin, R.A. 2019a. Data from: A mother's legacy: the strength of maternal effects in animal populations.

Moore, M.P., Whiteman, H.H. & Martin, R.A. 2019b. A mother's legacy: the strength of maternal effects in animal populations. *Ecol. Lett.* **22**:1620–1628.

Morrissey, M.B. & Wilson, A.J. 2010. pedantics: an r package for pedigree-based genetic simulation and pedigree manipulation, characterization and viewing. *Mol. Ecol. Resour.* **10**:711–719.

Mousseau, T. & Fox, C.W. 1998. *Maternal Effects as Adaptations*. Oxford University Press, Oxford.

Nakagawa, S., Ivimey-Cook, E.R., Grainger, M.J., O'Dea, R.E., Burke, S., Drobniak, S.M. *et al.* 2023. Method Reporting with Initials for Transparency (MeRIT) promotes more granularity and accountability for author contributions. *Nat. Commun.* **14**:1788, publisher: Nature Publishing Group.

Noble, D.W., Mcfarlane, S.E., Keogh, J.S. & Whiting, M.J. 2014. Maternal and additive genetic effects contribute to variation in offspring traits in a lizard. *Behav. Ecol.* **25**:633–640.

Ovaskainen, O., Cano, J.M. & Merilä, J. 2008. A Bayesian framework for comparative quantitative genetics. *Proc. R. Soc. B: Biol. Sci.* **275**:669–678, publisher: Royal Society.

Pick, J.L. 2024a. pedAgree: demographically explicit pedigree simulation. R package, version 0.0.1.

Pick, J.L. 2024b. squidSim: a flexible simulation tool for linear mixed models. R package, version 0.2.3.

Pick, J.L., Ebneter, C., Hutter, P. & Tschirren, B. 2016a. Disentangling genetic and prenatal maternal effects on offspring size and survival. *Am. Nat.* **188**:628–639.

Pick, J.L. & Hadfield, J.D. 2022. Data and code from: Decomposing phenotypic skew and its effects on the predicted response to strong selection.

Pick, J.L., Hutter, P. & Tschirren, B. 2016b. In search of genetic constraints limiting the evolution of egg size: direct and correlated responses to artificial selection on a prenatal maternal effector. *Heredity*. **116**:542–549.

Pick, J.L., Kasper, C., Allegue, H., Dingemanse, N.J., Dochtermann, N.A., Laskowski, K.L. *et al.* 2023. Describing posterior distributions of variance components: Problems and the use of null distributions to aid interpretation. *Methods Ecol. Evol.* **14**:2557–2574.

Pick, J.L., Lemon, H.E., Thomson, C.E. & Hadfield, J.D. 2022. Decomposing phenotypic skew and its effects on the predicted response to strong selection. *Nat. Ecol. & Evol.* **6**:774–785.

Pick, J.L., Nakagawa, S. & Noble, D.W. 2019a. Reproducible, flexible and high throughput

881 data extraction from primary literature: The metaDigitise R package. *Methods Ecol. Evol.*
882 **10**:426–431.

883 Pick, J.L., Postma, E. & Tschirren, B. 2019b. The more you get, the more you give: Positive
884 cascading effects shape the evolutionary potential of prenatal maternal investment. *Evol.*
885 *Lett.* **3**:412–423.

886 Postma, E. 2014. Four decades of estimating heritabilities in wild vertebrate populations:
887 Improved methods, more data, better estimates? In: *Quantitative Genetics in the Wild*
888 (A. Charmantier, D. Garant & L.E.B. Kruuk, eds.), pp. 16–33, Oxford University Press,
889 Oxford.

890 Pujol, B., Blanchet, S., Charmantier, A., Danchin, E., Facon, B., Marrot, P. *et al.* 2018. The
891 Missing Response to Selection in the Wild. *Trends Ecol. & Evol.* **33**:337–346.

892 Quéméré, E., Gaillard, J.M., Galan, M., Vanpé, C., David, I., Pellerin, M. *et al.* 2018. Between-
893 population differences in the genetic and maternal components of body mass in roe deer.
894 *BMC Evol. Biol.* **18**:39.

895 R Core Team. 2022. R: A Language and Environment for Statistical Computing. Place: Vienna,
896 Austria.

897 Regan, C.E., Pilkington, J.G., Bérénos, C., Pemberton, J.M., Smiseth, P.T. & Wilson, A.J.
898 2017. Accounting for female space sharing in St. Kilda Soay sheep (Ovis aries) results in
899 little change in heritability estimates. *J. Evol. Biol.* **30**:96–111.

900 Riska, B., Rutledge, J.J. & Atchley, W.R. 1985. Covariance between direct and maternal
901 genetic effects in mice, with a model of persistent environmental influences. *Genet. research*
902 **45**:287–297.

903 Roach, D.A. & Wulff, R.D. 1987. Maternal effects In plants. *Annu. Rev. Ecol. Syst.* **18**:209–
904 235, iSBN: 00664162.

Räsänen, K. & Kruuk, L.E.B. 2007. Maternal effects and evolution at ecological time-scales. *Funct. Ecol.* **21**:408–421.

Réale, D., Martin, J., Coltman, D.W., Poissant, J. & Festa-Bianchet, M. 2009. Male personality, life-history strategies and reproductive success in a promiscuous mammal. *J. Evol. Biol.* **22**:1599–1607.

Satoh, M., Hicks, C., Ishii, K. & Furukawa, T. 2002. Choice of statistical model for estimating genetic parameters using restricted maximum likelihood in swine. *J. Animal Breed. Genet.* **119**:285–296.

Thompson, R. 1976. The Estimation of Maternal Genetic Variances. *Biometrics*. **32**:903.

Thomson, C.E., Bayer, F., Cassinello, M., Crouch, N., Heap, E., Mittell, E. *et al.* 2017. Selection on parental performance opposes selection for larger body size in a wild population of blue tits. *Evolution*. **71**:716–732.

Varona, L., Munilla, S., Casellas, J., Moreno, C. & Altarriba, J. 2015. Consequences of paternally inherited effects on the genetic evaluation of maternal effects. *Genet. Sel. Evol.* **47**:63.

Walling, C.A., Stamper, C.E., Smiseth, P.T. & Moore, A.J. 2008. The quantitative genetics of sex differences in parenting. *Proc. Natl. Acad. Sci. United States Am.* **105**:18430–18435.

Willham, R.L. 1963. The covariance between relatives for characters composed of components contributed by related individuals. *Biometrics*. **19**:18–27.

Willham, R.L. 1972. The role of maternal effects in animal breeding: III. Biometrical aspects of maternal effects in animals. *J. Animal Sci.* **35**:1288–1293.

Wilson, A., Kruuk, L. & Coltman, D. 2005a. Ontogenetic Patterns in Heritable Variation for

927 Body Size: Using Random Regression Models in a Wild Ungulate Population. *Am. Nat.*
928 **166**:E177–E192.

929 Wilson, A.J., Coltman, D.W., Pemberton, J.M., Overall, A.D.J., Byrne, K.A. & Kruuk, L.E.B.
930 2005b. Maternal genetic effects set the potential for evolution in a free-living vertebrate
931 population. *J. Evol. Biol.* **18**:405–414.

932 Wilson, A.J. & Reale, D. 2006. Ontogeny of additive and maternal genetic effects: lessons
933 from domestic mammals. *Am. Nat.* **167**:E23–E38.

934 Wolf, J., Brodie III, E., Cheverud, J., Moore, A.J. & Wade, M. 1998. Evolutionary conse-
935 quences of indirect genetic effects. *Trends Ecol. Evol.* **13**:64–69.

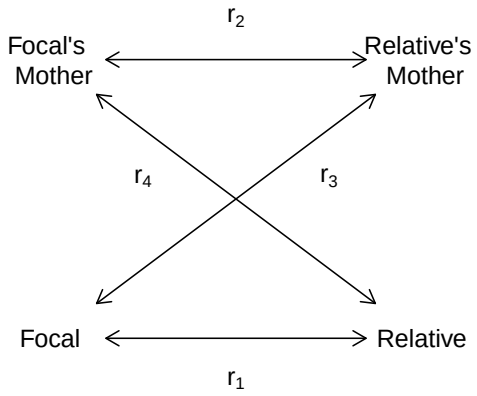
936 Young, E.A. & Postma, E. 2023a. Low interspecific variation and no phylogenetic signal in
937 additive genetic variance in wild bird and mammal populations. *Ecol. Evol.* **13**:e10693.

938 Young, E.A. & Postma, E. 2023b. Replication Data for: Low interspecific variation and no
939 phylogenetic signal in additive genetic variance in wild bird and mammal populations.

Supplementary Material

S1 Expected phenotypic covariance between relatives

The expected phenotypic covariance of two related individual in the presence of maternal genetic variation is determined not only by V_A, but also potentially by V_{Mg} and $COV_{A,Mg}$, depending on how the two individual's mothers are related, and how related the two individuals are to the other individual's mother. Figure S1 shows how the expected phenotypic covariance can be derived. Note that these phenotypic covariance may be further affected by other factors such as dominance, that we do not include here.



$$r_1 V_A + r_2 V_{Mg} + (r_3 + r_4) COV_{A,Mg}$$

Figure S1: Calculation of covariance between related individuals. The arrows represent the relevant relatedness for the calculation of the covariance and the equation below shows how these are added to give the covariance.

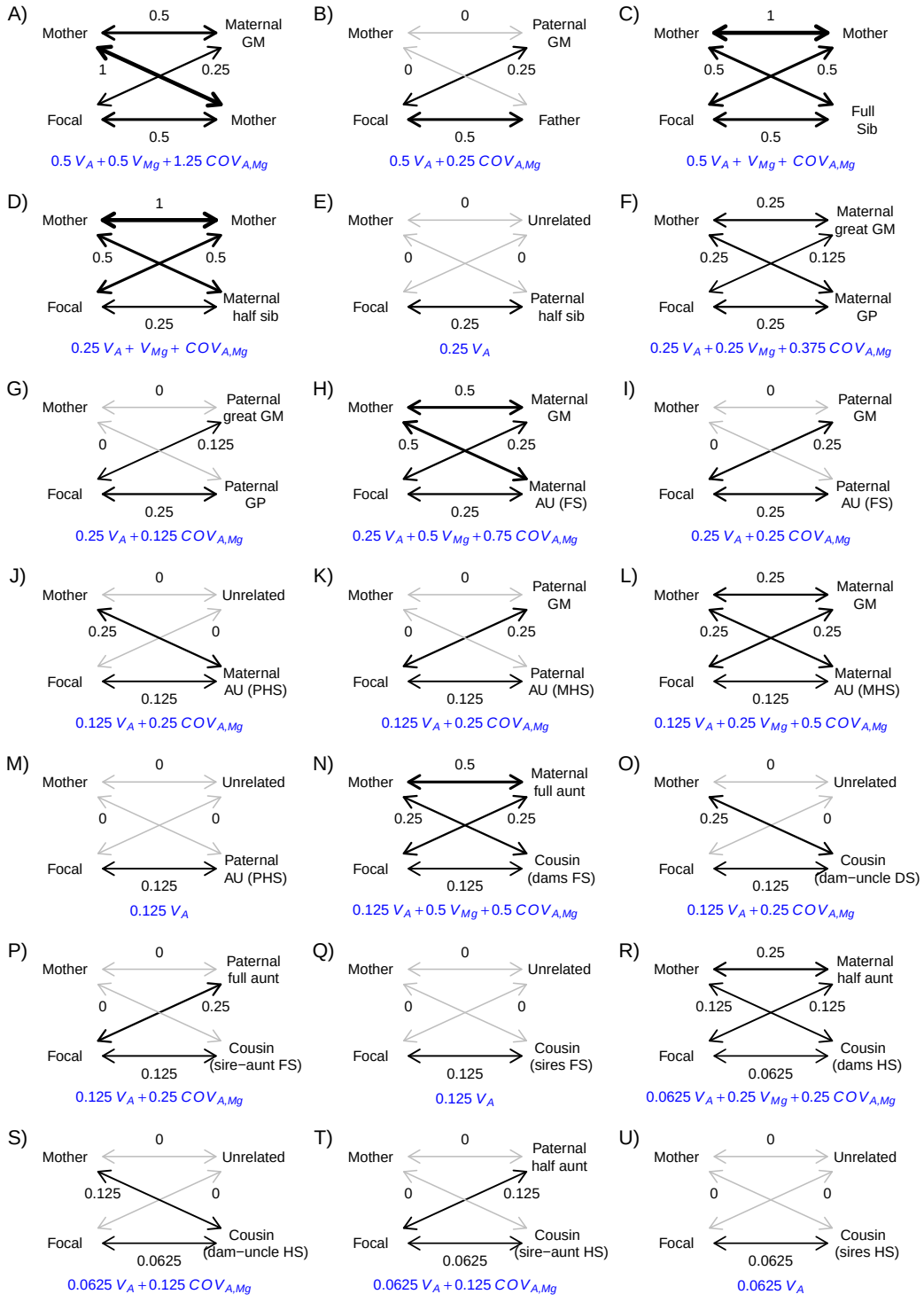


Figure S2: Calculation of the expected phenotypic covariance for the relationships shown in Table 2. Each panel shows how total expected covariance between a focal individual (bottom left) and a given relative (bottom right is calculated from the relatedness between the two relatives and their mothers (top left and right). The arrows represent the relevant relatedness for the calculation of the covariance and their weighting represent how related the two individuals are.

S2 Estimates from the literature

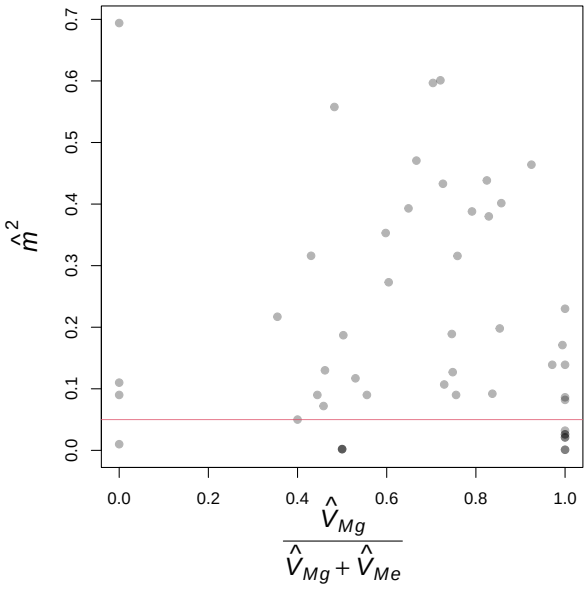


Figure S3: Relationship between $\hat{m}^2$ and the proportion of $\hat{V}_M$ due to$\hat{V}_{Mg}$, illustrating the justification for excluding the estimates with $\hat{m}^2 > 0.05$. Red lines shows the 0.05 cut off, below which all estimates of the proportion are either exactly 0, 0.5 or 1.

Table S1: Estimates of maternal genetic variation from previous studies. A csv of this table is included in the data accompanying the paper.

Study	Species	Population	Trait	Age	h_1^2	c_1^2	h_2^2	m_2^2	c_2^2	Source
1	Soay Sheep	wild	Birth weight	Juvenile	0.091	0.201	0.079	0.077	0.140	T S4 and S5
1	Soay Sheep	wild	Lamb Foreleg	Juvenile	0.155	0.063	0.140	0.033	0.039	T S4 and S5
1	Soay Sheep	wild	Lamb Hindleg	Juvenile	0.196	0.068	0.161	0.068	0.022	T S4 and S5
1	Soay Sheep	wild	Lamb Weight	Juvenile	0.116	0.100	0.066	0.095	0.032	T S4 and S5
1	Soay Sheep	wild	Lamb Metacarpal	Juvenile	0.509	0.081	0.402	0.139	0.000	T S4 and S5
1	Soay Sheep	wild	Lamb Jaw	Juvenile	0.303	0.145	0.203	0.141	0.048	T S4 and S5
1	Soay Sheep	wild	Yearling Foreleg	Adult	0.157	0.092	0.108	0.078	0.029	T S4 and S5
1	Soay Sheep	wild	Yearling Hindleg	Adult	0.307	0.168	0.271	0.094	0.093	T S4 and S5
1	Soay Sheep	wild	Yearling Weight	Adult	0.190	0.099	0.156	0.062	0.055	T S4 and S5
1	Soay Sheep	wild	Yearling Metacarpal	Adult	0.618	0.000	0.618	0.000	0.000	T S4 and S5
1	Soay Sheep	wild	Yearling Jaw	Adult	0.672	0.089	0.672	0.000	0.090	T S4 and S5
1	Soay Sheep	wild	Adult Foreleg	Adult	0.296	0.005	0.286	0.021	0.000	T S4 and S5
1	Soay Sheep	wild	Adult Hindleg	Adult	0.458	0.063	0.426	0.077	0.015	T S4 and S5
1	Soay Sheep	wild	Adult Weight	Adult	0.273	0.057	0.246	0.082	0.000	T S4 and S5
1	Soay Sheep	wild	Adult Metacarpal	Adult	0.631	0.018	0.610	0.032	0.000	T S4 and S5
1	Soay Sheep	wild	Adult Jaw	Adult	0.677	0.010	0.677	0.000	0.010	T S4 and S5
2	Soay Sheep	wild	Birth date	Juvenile			0.070	0.690		T 1
2	Soay Sheep	wild	Birth weight	Juvenile			0.160	0.250		T 1
2	Soay Sheep	wild	Lamb Foreleg	Juvenile				0.130		T 1
2	Soay Sheep	wild	Lamb Hindleg	Juvenile				0.140		T 1
2	Soay Sheep	wild	Lamb weight	Juvenile				0.200		T 1
3	Soay Sheep	wild	Birth weight	Juvenile			0.014	0.169	0.029	T 2
3	Soay Sheep	wild	Birth date	Juvenile			0.032	0.255	0.138	T 2
3	Soay Sheep	wild	August weight	Juvenile			0.036	0.135	0.004	T 2
4	Soay Sheep	wild	Birth weight	Juvenile	0.116	0.214	0.114	0.165	0.108	T 1
4	Soay Sheep	wild	Birth date	Juvenile	0.127	0.316	0.072	0.315	0.065	T 1
4	Soay Sheep	wild	Natal litter size	Juvenile	0.077	0.252	0.109	0.211	0.142	T 1

5	Red Squirrels	wild	Female LRS	Adult	0.001	0.080	0.001	0.020	0.030	T 1 in both
5	Red Squirrels	wild	Lifespan	Adult			0.001	0.040	0.050	T 1
5	Red Squirrels	wild	Mean ARS	Adult			0.001	0.050	0.040	T 1
5	Red Squirrels	wild	AFB	Adult			0.000	0.060	0.070	T 1
6	Roe Deer	wild	Juvenile body mass	Juvenile	0.100	0.120	0.050	0.000	0.110	T 1 and 2
6	Roe Deer	wild	Juvenile body mass	Juvenile	0.530	0.080	0.440	0.230	0.000	T 1 and 2
7	Red Deer	wild	Birth weight	Juvenile	0.134	0.370	0.041	0.429	0.035	F 5
8	Red Deer	wild	Birth weight	Juvenile	0.221	0.353	0.177	0.307	0.081	T 2
8	Red Deer	wild	Birth leg	Juvenile	0.380	0.169	0.335	0.170	0.001	T 2
8	Red Deer	wild	Neonatal survival	Juvenile	0.038	0.000	0.022	0.021	0.000	T 2
8	Red Deer	wild	Survival age 1	Juvenile	0.063	0.032	0.051	0.026	0.000	T 2
8	Red Deer	wild	Survival age 2	Adult	0.047	0.031	0.052	0.026	0.000	T 2
8	Red Deer	wild	Female AFR	Adult	0.164	0.001	0.001	0.001	0.001	T 2
8	Red Deer	wild	Female ABS	Adult	0.040	0.000	0.040	0.000	0.000	T 2
8	Red Deer	wild	Male ABS	Adult	0.014	0.000	0.018	0.000	0.000	T 2
8	Red Deer	wild	Adult longevity	Adult	0.188	0.001	0.130	0.001	0.001	T 2
8	Red Deer	wild	Jaw	Adult	0.500	0.000	0.447	0.001	0.000	T 2
8	Red Deer	wild	Endocranial volume	Adult	0.776	0.001	0.629	0.001	0.000	T 2
8	Red Deer	wild	Leg	Adult	0.583	0.001	0.502	0.001	0.001	T 2
9	Red Deer	wild	Anti-Tc IgA	Juvenile			0.033	0.420	0.177	SM
9	Red Deer	wild	Total IgA	Juvenile			0.042	0.361	0.077	SM
9	Red Deer	wild	Anti-Tc IgM	Juvenile			0.059	0.314	0.157	SM
9	Red Deer	wild	Total IgM	Juvenile			0.036	0.314	0.119	SM
9	Red Deer	wild	Anti-Tc IgG	Juvenile			0.019	0.269	0.288	SM
9	Red Deer	wild	Total IgG	Juvenile			0.000	0.086	0.000	SM
10	Bighorn Sheep	wild	June weight age 0	Juvenile			0.000	0.197		T 3
10	Bighorn Sheep	wild	June weight age 1	Adult			0.447	0.135		T 3
11	Bighorn Sheep	wild	Boldness	Adult			0.390	0.000		T 1
12	American Bellflower	breeding design	Seed mass	Adult			0.064	0.000	0.694	T 1
12	American Bellflower	breeding design	Days to germination	Adult			0.361	0.303		T 1
12	American Bellflower	breeding design	Rosette size	Adult			0.243	0.260		T 1

12	American Bellflower	breeding design	Days to flower	Adult			0.426	0.433	0.168	T 1
12	American Bellflower	breeding design	Biomass	Adult			0.368	0.133		T 1
13	Squirrel Monkeys	managed	Female neonate mass	Juvenile	0.190	0.318	0.092	0.344	0.057	T 2
13	Squirrel Monkeys	managed	Male neonate mass	Juvenile	0.159	0.256	0.121	0.240	0.076	T 2
14	Cuvier's gazelle	managed	Juvenile survival	Juvenile	0.067	0.247	0.115	0.136	0.180	T 2

1: Bérénos *et al.* 2014 , 2:Beraldi *et al.* 2007 , 3:Regan *et al.* 2017 , 4:Wilson *et al.* 2005b , 5:McFarlane *et al.* 2014, 2015 , 6:Quéméré *et al.* 2018 , 7:Kruuk & Hadfield 2007 , 8:Gauzere *et al.* 2020b , 9:Gauzere *et al.* 2021 , 10:Wilson *et al.* 2005a , 11:Réale *et al.* 2009 , 12:Galloway *et al.* 2009 , 13:Blomquist & Williams 2013 , 14:Ibáñez *et al.* 2014

S3 Non-sibling maternal links

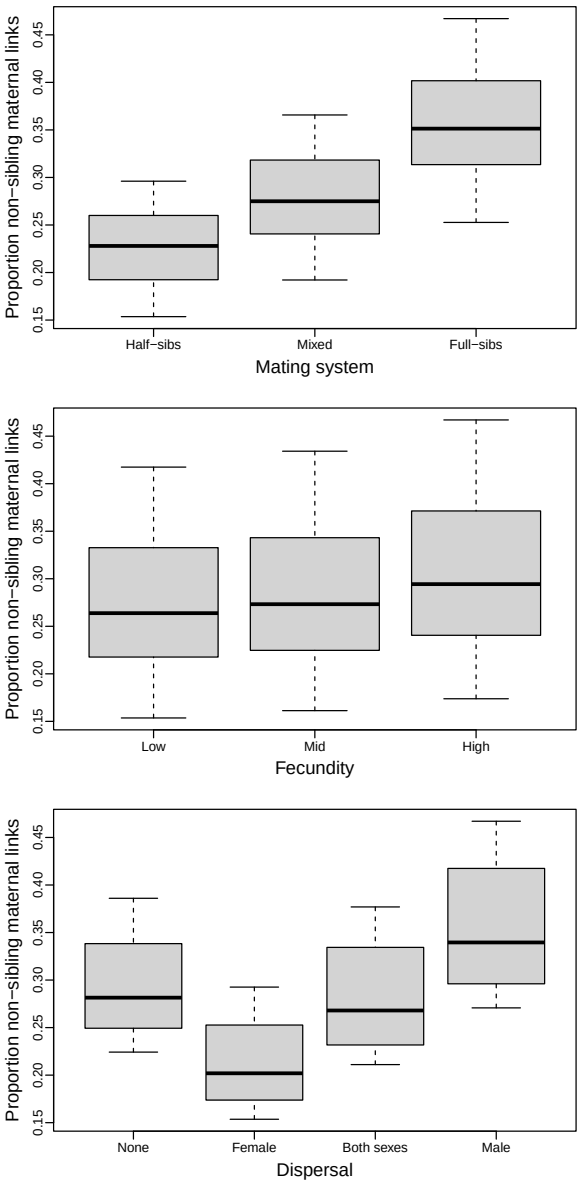


Figure S4: Variation in proportion of non-sibling maternal links across 36 pedigree types.

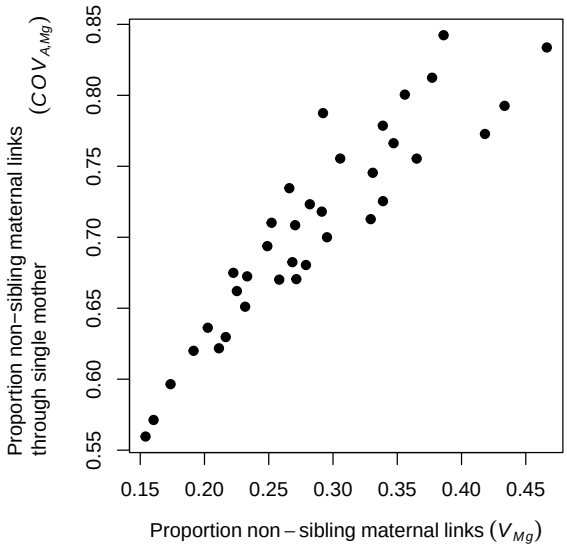


Figure S5: Relationship between proportion of non-sibling links informative for V_{Mg} and $COV_{A,Mg}$ across 36 pedigree types.

S4 Effect of Pedigree depth on non-sibling maternal links

We examined the effect of pedigree depth on the the proportion of non-sibling maternal links. Pedigree depth has previously been shown to have a strong effect on the estimation of quantitative genetic parameters (Kruuk & Hadfield, 2007). In this case it may affect the build up of non-sibling maternal links in the pedigree. To examine this, we simulated pedigrees that varied in the number of discrete generations, from 2 to 10, across the four immigration rates described in the main text. For each pedigree we simulated 100 females per generation. We opted to have different sample sizes across the different pedigrees, rather than varying the number of females per generation. All pedigrees were simulated with the intermediate fecundity (6 offspring per female) and mating system (probability of 0.75 that paired male will sire offspring) parameters used in the main simulations. 50 pedigrees were simulated per immigration rate and pedigree depth combination. These simulations showed that proportion of non-sibling maternal links was reduced in very shallow pedigrees, but stabilised after a couple of generations (Figure S6).

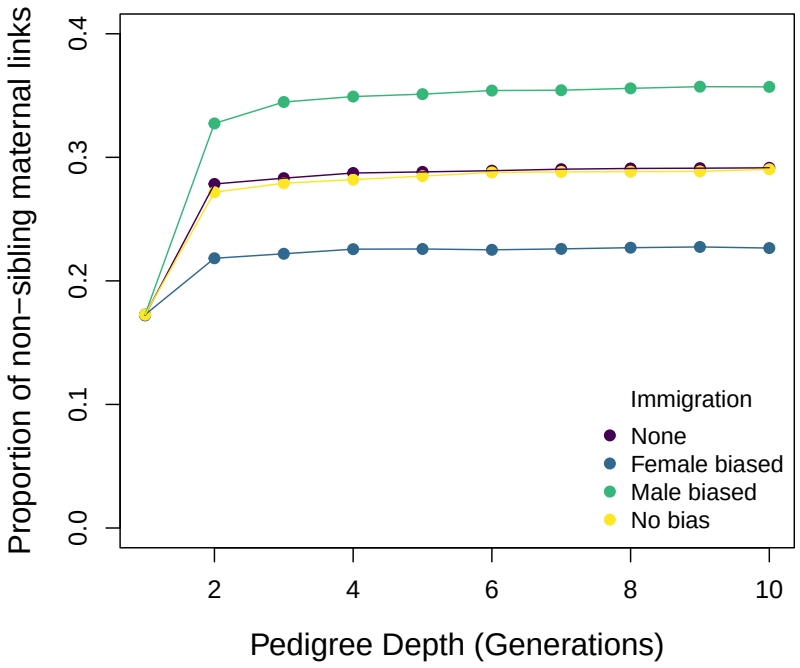


Figure S6: The effect of pedigree depth on the proportion of non-sibling maternal links, across the different immigration scenarios.

S5 Results from scenarios K and L

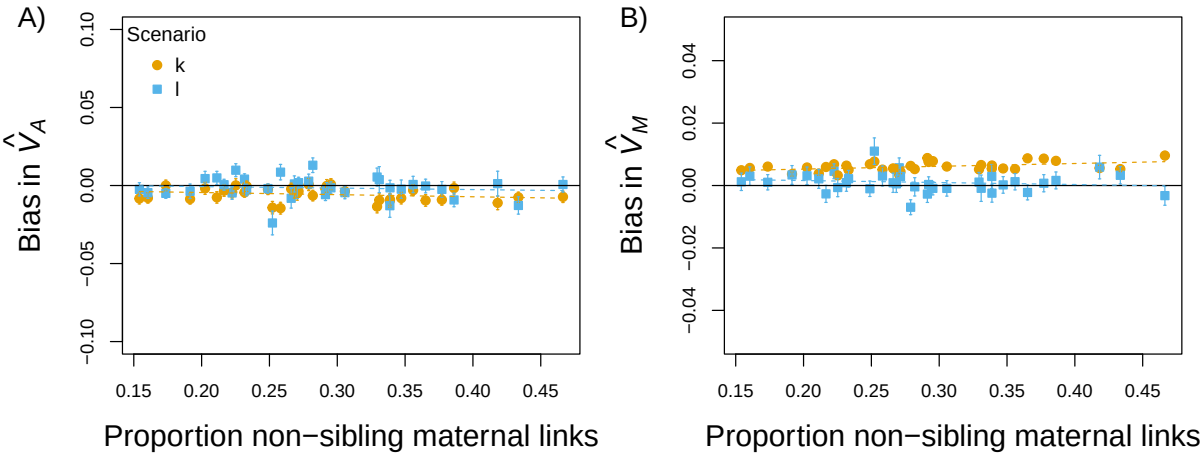


Figure S7: Bias in $\hat{V}_A$ (A) and $\hat{V}_M$ (B) in relation to the proportion of non-sibling maternal links across two simulated scenarios (indicated by the colours; see Table 3) and pedigree structures. Both scenarios were simulated with V_A and no V_{Mg}. Dotted lines are predictions from a simple linear model, the purpose of which is just to help illustrate the pattern. Error bars show the standard error across simulations. Note that for some simulations the errors bars are too small to see.

S6 Sampling covariance and risk of bias

Within each statistical model, we can look the estimated sampling covariance between two parameters, which gives us information about how well the model is able to independently estimate the two parameters, also known as the identifiability of the parameters. The sampling covariance between the $\hat{V}_A$ and $\hat{V}_{Mc}$ therefore tells us how well the model is able to separate the two; if the covariance is strongly negative, then the model is struggling to tell where the variance is coming from. We might expect when there is more risk of bias from unmodelled V_{Mg} that the covariance is larger.

From looking to the sampling variance from our simulations, this indeed this appears to be the case. Figure S8 shows a clear relationship between the proportion of non-sibling maternal links and the estimated sampling covariance, with the sampling covariance becoming more negative as the proportion increased.

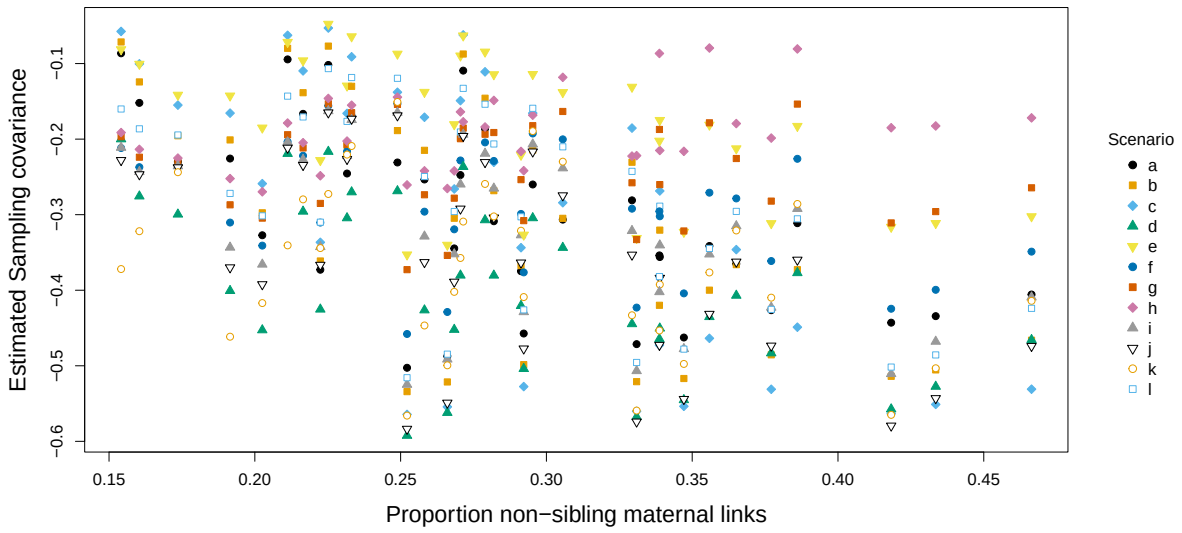


Figure S8: Relationships between of pedigree structure (in terms of non-sibling maternal links) and sampling covariance between $\hat{V}_A$ and $\hat{V}_{Mc}$

We can also look more closely at how the sampling covariance varied across the different pedigrees simulated. Figure S9 shows that the clearest effect on the sampling covariance is due the mating system, in other words the amount of half siblings. This makes sense as generally maternal variance is harder to separate from V_A when siblings share both the same

parents and the same environment.

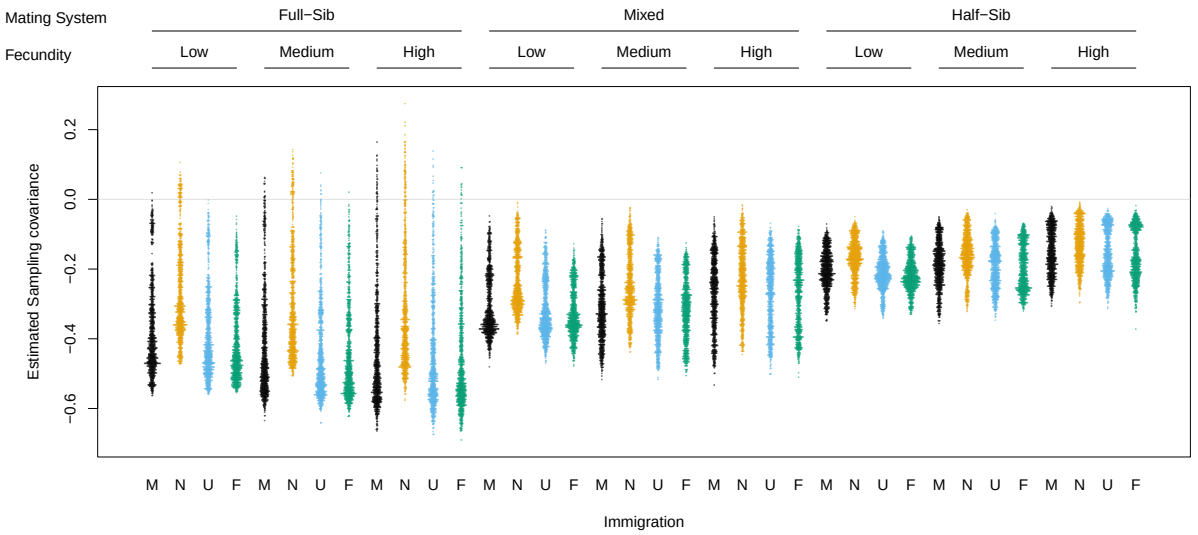


Figure S9: Variation in the sampling covariance between V_A and $\hat{V}_{Mc}$ across different simulated pedigree structures. M, N, U and F, stand for male-biased, no, unbiased and female-biased immigration respectively.

It would be interesting to know whether the sampling covariance varied systematically across the simulated scenarios, and so whether considering the sampling covariance would enable us to tell something about the presence of unmodelled V_{Mg} in a real datasets. However, the sampling covariance was not clearly consistently affected by the actual underlying parameters. Figure S10 shows that when comparing the first three scenarios (no V_A, varying levels of V_{Me} and V_{Mg}) the range of sampling covariances was similar, and in fact the only scenario without V_{Mg} (scenario C), had the widest range of sampling covariance. The sampling covariance appeared to be affected by the presence of $COV_{A,Mg}$ (G, H, I and J), although again the range of these covariances was almost entirely covered by the scenario without V_{Mg}. Positive sampling covariances were rare, and interestingly only occurred when $COV_{A,Mg}$ was strong and positive (scenario H).

In conclusion, the sampling covariance largely indicates the ability of the model to separate $\hat{V}_A$ and $\hat{V}_{Mc}$, and so is strongly affects by factors such as the number of half siblings. The sampling covariance relates in some way to the risk of bias in V_A, but is not affected is a clear way by the actual underlying parameters, and so cannot be used to assess whether the

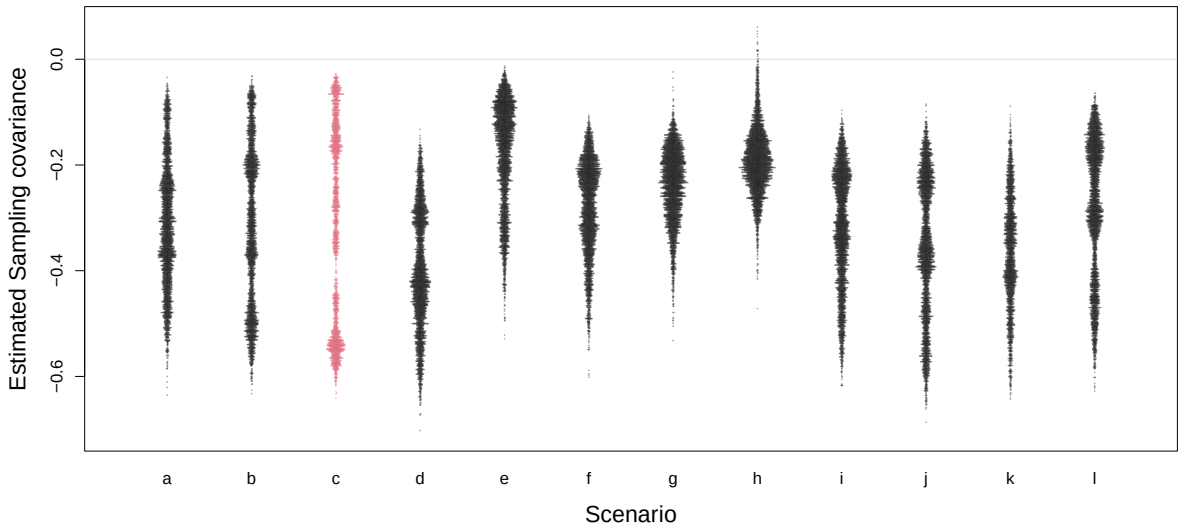


Figure S10: Variation in the sampling covariance between $\hat{V}_A$ and $\hat{V}_{Mc}$ across different simulated scenarios. Red points indicate the scenario where V_{Mg} was 0.

estimates are likely to actually be biased.

S7 Small pedigree simulations

In the main text, only the results from a single pedigree (small pedigree with unbiased immigration) are shown. The results across all pedigree types were very highly correlated:

Table S2: Correlation between mean absolute error across scenarios in different pedigree structures.

	nl small	fl small	ml small	ul small	nl medium	fl medium	ml medium	ul medium
nl small	1							
fl small	0.966	1						
ml small	0.973	0.976	1					
ul small	0.976	0.978	0.983	1				
nl medium	0.941	0.927	0.936	0.946	1			
fl medium	0.952	0.978	0.972	0.98	0.954	1		
ml medium	0.943	0.956	0.961	0.967	0.976	0.984	1	
ul medium	0.949	0.974	0.972	0.975	0.969	0.992	0.991	1

Figures S11-S13 show the bias, precision and accuracy of $\hat{V}_A$, $\hat{V}_M$ and $\hat{V}_{A_t}$ across all scenarios and pedigree structures.

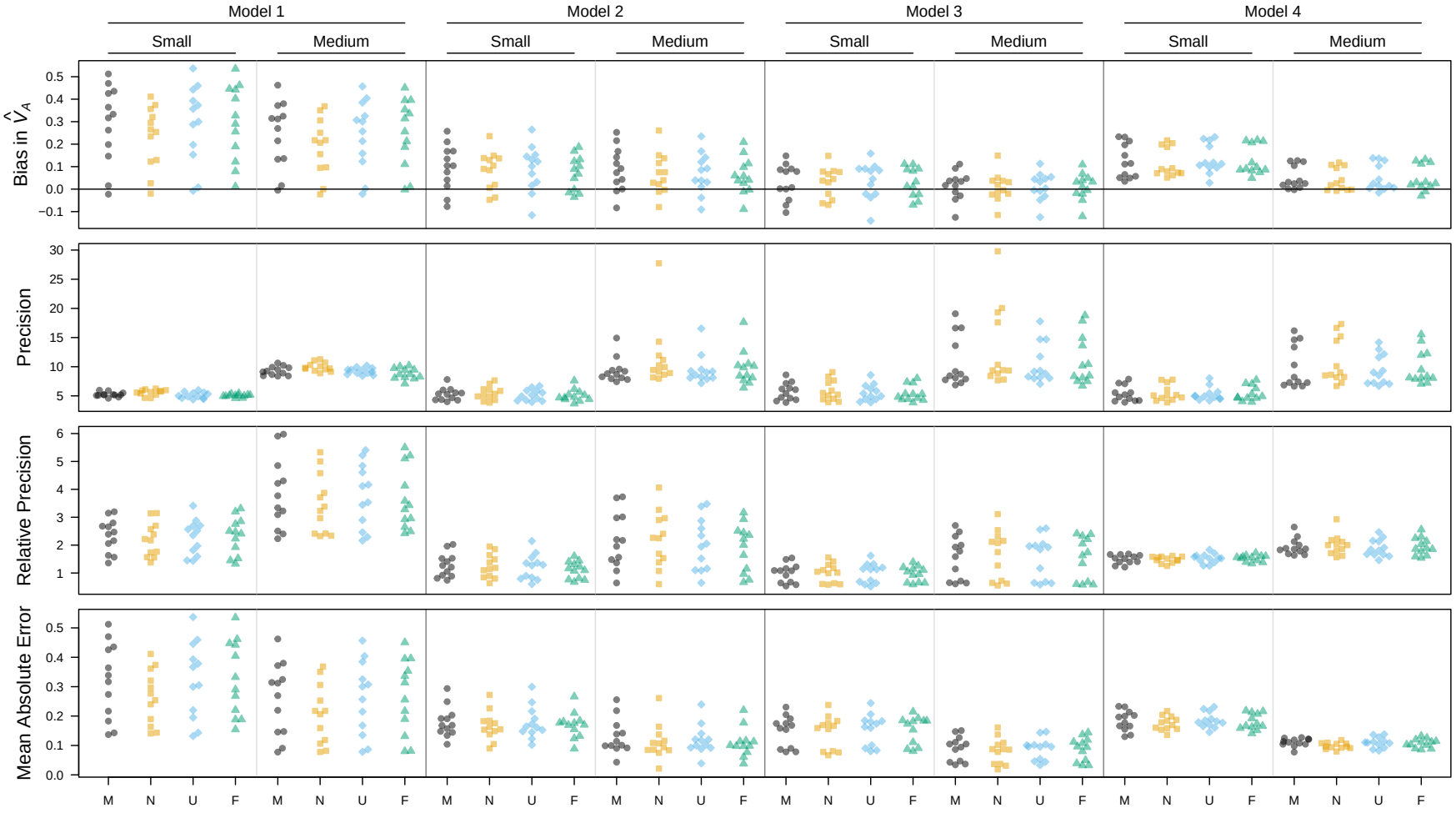


Figure S11: Bias, precision, relative precision and accuracy (measured as absolute mean error) in $\hat{V}_A$ from 4 different models. from simulations of small and medium sized pedigrees with varying immigration rates. Points show the metrics calculated for each of the 12 scenarios (see Table 3). M, N, U and F, stand for male-biased, no, unbiased and female-biased immigration respectively.

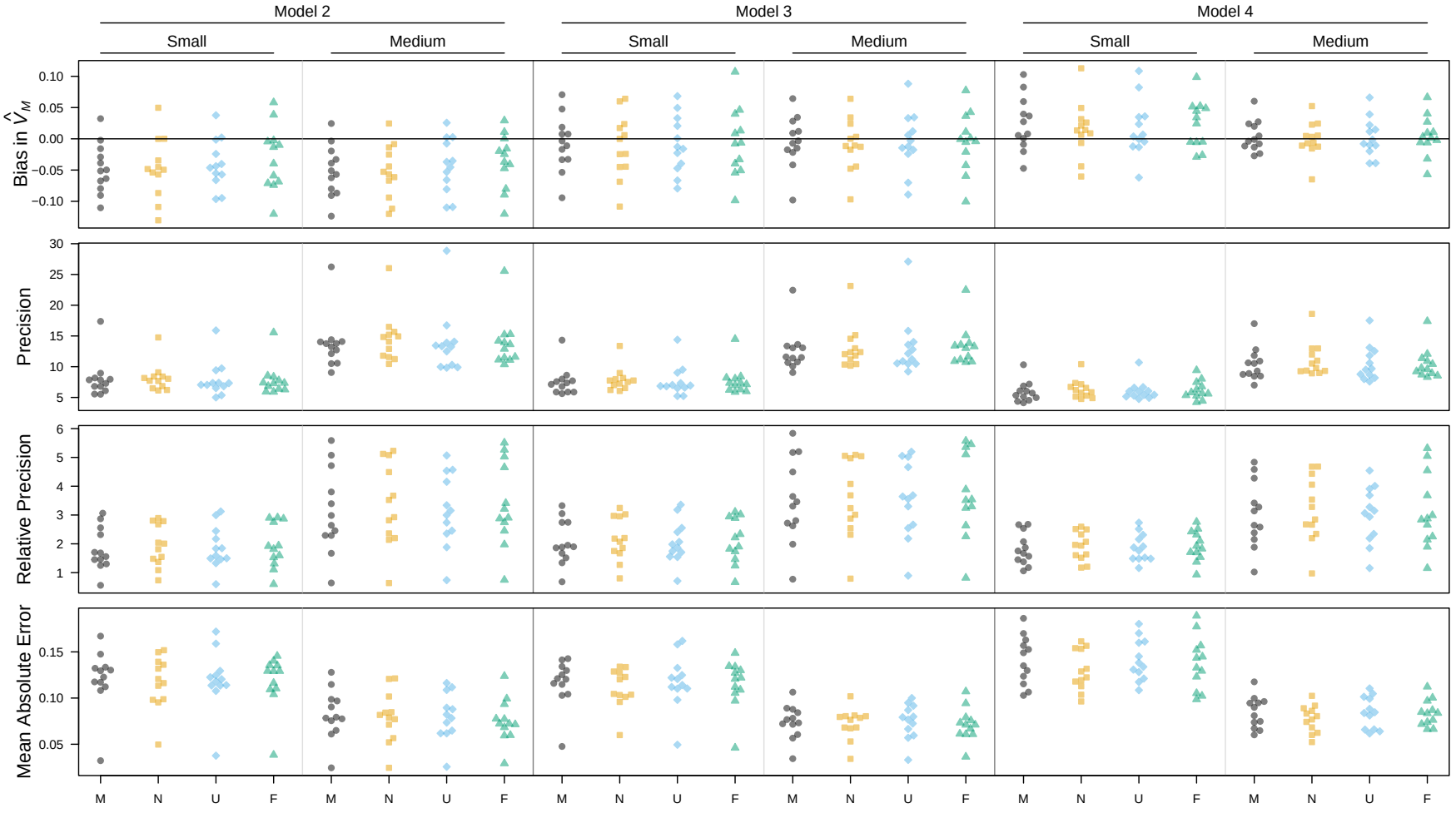


Figure S12: Bias, precision, relative precision and accuracy (measured as absolute mean error) in $\hat{V}_M$ from 3 different models. from simulations of small and medium sized pedigrees with varying immigration rates. Points show the metrics calculated for each of the 12 scenarios (see Table 3). M, N, U and F, stand for male-biased, no, unbiased and female-biased immigration respectively.

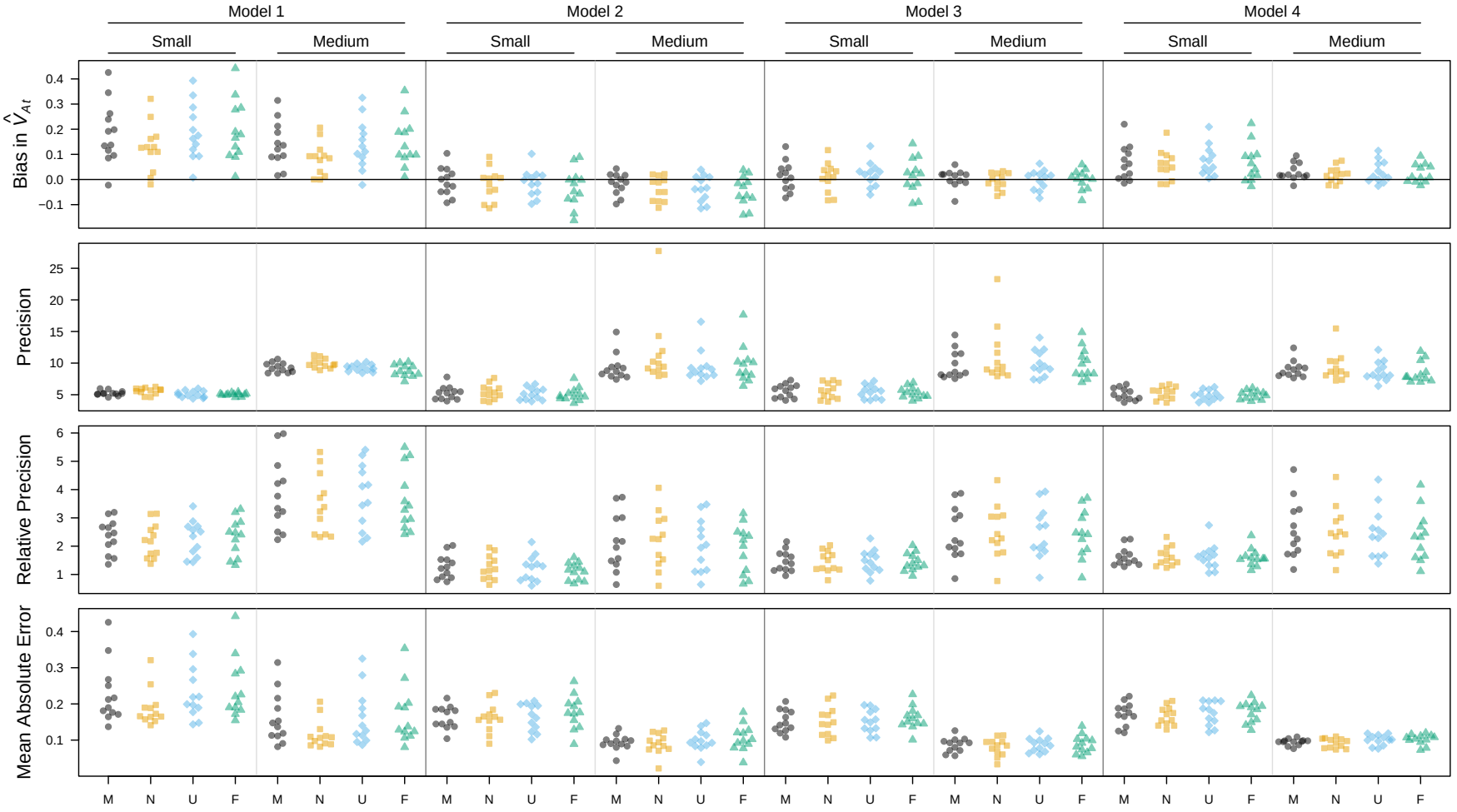


Figure S13: Bias, precision, relative precision and accuracy (measured as absolute mean error) in $\hat{V}_{A_t}$ from 4 different models. from simulations of small and medium sized pedigrees with varying immigration rates. Points show the metrics calculated for each of the 12 scenarios (see Table 3). M, N, U and F, stand for male-biased, no, unbiased and female-biased immigration respectively.

S7.1 Estimates of $\hat{V}_{Mg}$ and $\hat{COV}_{A,Mg}$ in small pedigrees

In the main text we focus primarily on $\hat{V}_A$, $\hat{V}_M$ and $\hat{V}_{A_t}$, and how running more complex models affects the estimation of these parameters. As noted in the main text, $\hat{V}_A$ was less accurate in model 4 than model 3, in scenarios where there was no underlying V_A (Figure 7). This was predominantly due to a larger upward bias in $\hat{V}_A$ in model 4. $\hat{V}_A$ also showed an upward bias in model 4 in scenarios where there was no V_{Mg}.

In Figure S14, we additionally consider $\hat{V}_{Mg}$ and $\hat{COV}_{A,Mg}$. As with $\hat{V}_A$, $\hat{V}_{Mg}$ was also more upwardly biased in scenarios with no V_{Mg} in model 4 than model 3, a bias which was increased at small sample sizes. In many scenarios, $\hat{COV}_{A,Mg}$ estimated in model 4 was downwardly biased. This was particularly pronounced in scenarios where there was no V_A or V_{Mg} simulated, and so no $COV_{A,Mg}$, and again more so in small pedigrees. When measured as a correlation ($\hat{r}_{A,Mg}$), we can see that this was quite a substantial bias, estimating moderate negative correlations when none was simulated.

We suspect that this is a non-identifiability issue. In the presence of V_{Mg} and $COV_{A,Mg}$, the phenotypic variance is $V_P = V_A + V_{Mg} + V_{Me} + COV_{A,Mg} + V_\epsilon$ (Willham, 1972). This means that the same phenotypic variance can be described by low variances and no covariance, or high variances and a negative covariance. Indeed, when there is a negative $COV_{A,Mg}$, and we do not model it, the model compensates by underestimating $\hat{V}_A$ and $\hat{V}_{Mg}$ (Figure S14).

At small sample sizes, the model clearly struggled to distinguish between these scenarios. As the variances cannot be negative, this causes this uncertainty to be biased in one direction and as a result the model overestimated the variances and estimated a negative covariance.

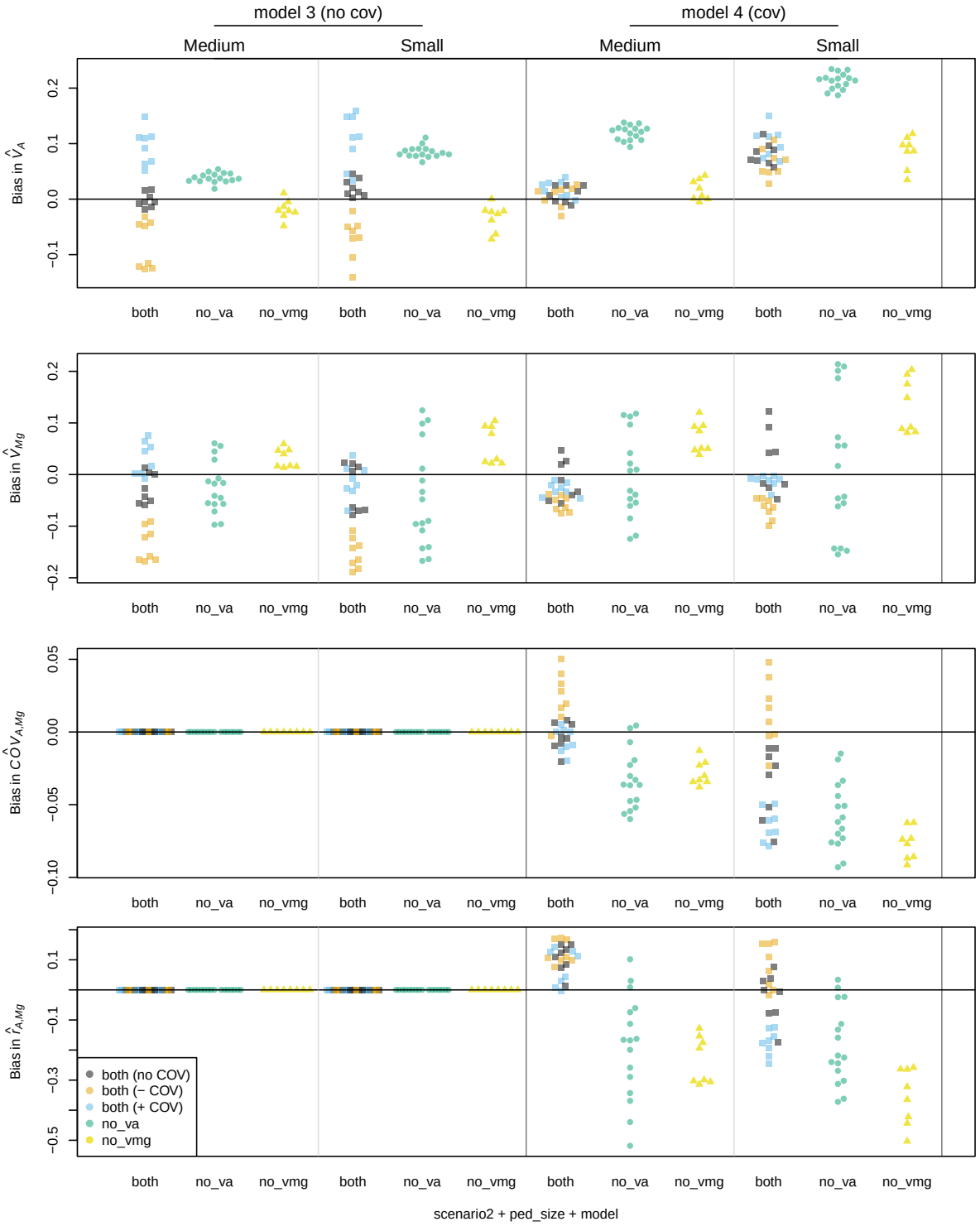


Figure S14: Bias in $\hat{V}_A$, $\hat{V}_{Mg}$, $\hat{COV}_{A,Mg}$ and $\hat{r}_{A,Mg}$ from small and medium sized pedigree simulations. Scenarios are separated on the x axis by whether V_A and V_{Mg} were simulated.

S7.2 Root Mean Squared Error

Root mean squared error (RMSE) is another commonly used metric to assess accuracy (or a measure of average directionless deviation from the true value). RMSE and mean absolute error (MAE, as use din the main text) are very similar.

In this case the use of either metric makes no difference to the inference. The correlation between the two metrics, measured across $\hat{V}_A$, $\hat{V}_M$ and $\hat{V}_{A_t}$ is very high (¿0.98; Figure S15).

If we compare the figure below (Figure S16) with Figure 7 in the main text, we can see the results are almost identical, and the inferences we would make are the same.

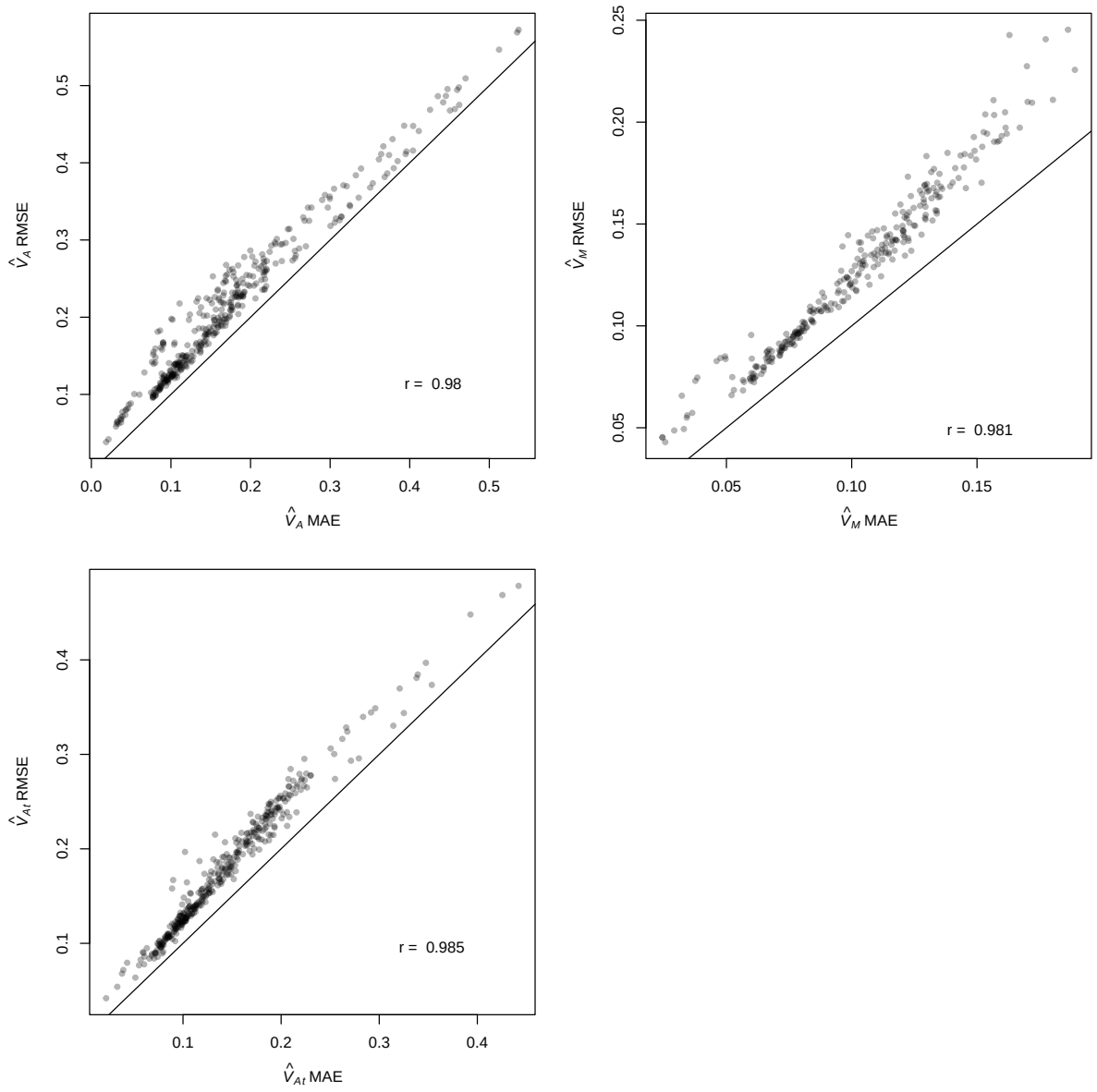


Figure S15: Comparison between root mean squared error (RMSE) and mean absolute error (MAE)

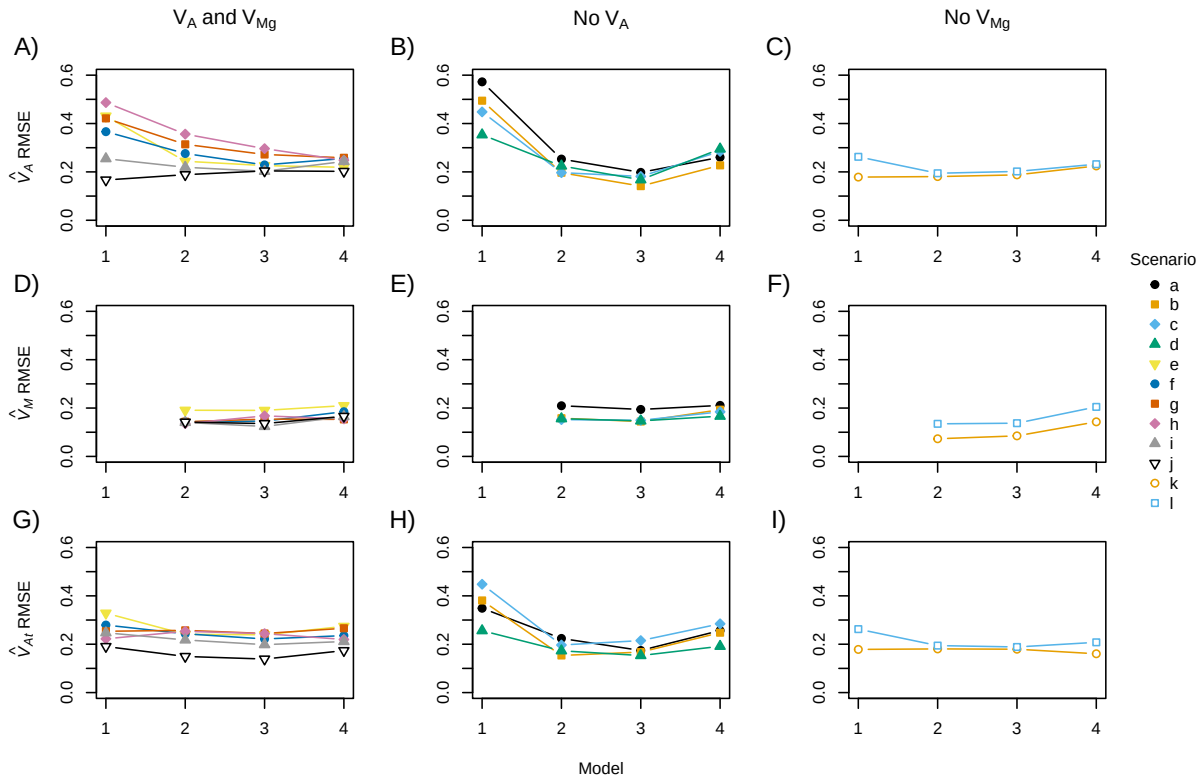


Figure S16: Accuracy (measured as mean root mean squared error (RMSE)) of $\hat{V}_A$ (A, B and C) $\hat{V}_M$ (D, E and F) and $\hat{V}_{A_t}$ (G, H and I), estimated using 4 different models: 1) estimated $\hat{V}_A$ only, 2) estimated $\hat{V}_A$ and $\hat{V}_{Mc}$, 3) estimated $\hat{V}_A$, $\hat{V}_{Mg}$ and $\hat{V}_{Me}$ and 4) estimated $\hat{V}_A$, $\hat{V}_{Mg}$, $\hat{V}_{Me}$ and $C\hat{O}V_{A,Mg}$. Data from these plots is from simulations of small pedigrees with unbiased immigration, over 12 scenarios (shown with different symbols and colours), which are separate out across the three columns, to show scenarios with both V_A and V_{Mg}, no V_A and no V_{Mg}, respectively.

S7.3 Model Convergence

Across all pedigrees, models 1-3 had no problems running in ASReml (i.e. occasional warnings about singularities etc, but no errors). Only the models specifying a covariance had any problems running. These models gave a convergence error in 8.65% of models. This was not equally distributed across scenarios; scenarios with no simulated V_A (scenarios 1-4) often did not converge (15-25%), compared to under 5% in most other scenarios (Figure S17).

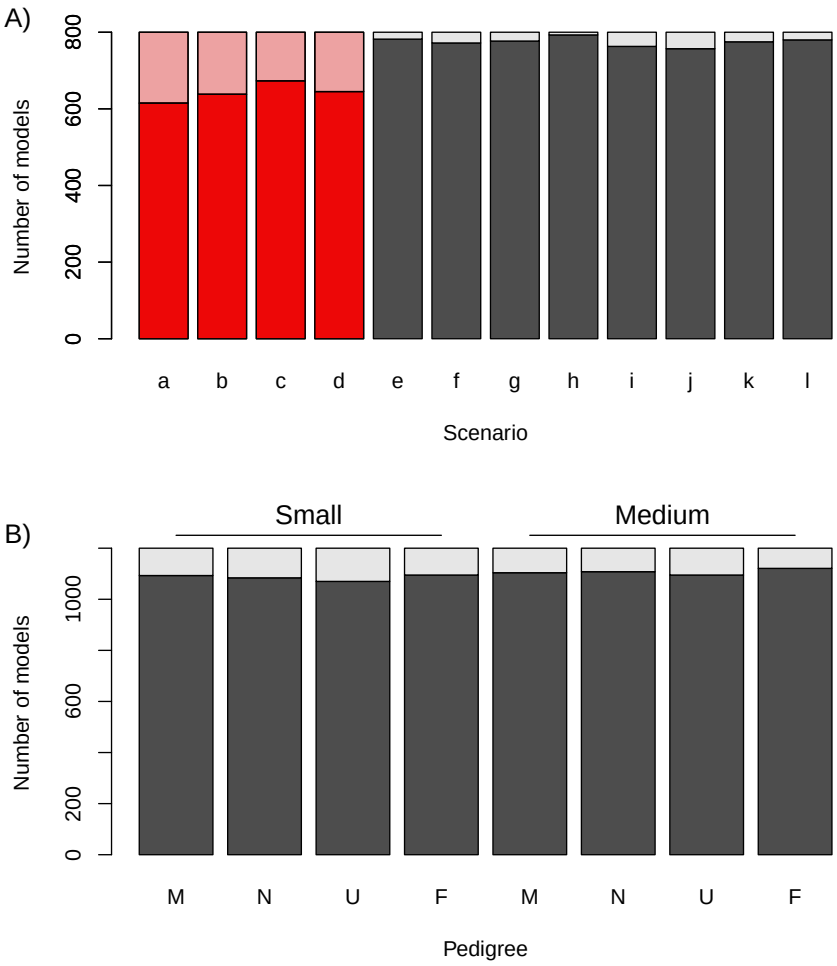


Figure S17: Number of datasets for which model 4 converged (dark) and did not converge(light), split by scenario (A) and pedigree type (B). M, N, U and F, stand for male-biased, no, unbiased and female-biased immigration respectively. In A), scenarios 1-4 are highlighted in red as these are the scenarios where no V_A was simulated.

These are also the same scenarios in which model 3 clearly performs better than model 4 (Figure S11). We wanted to make sure that the difference between these models wasn't

caused by including model 3 estimates from datasets where model 4 did not converge. The mean absolute error in $\hat{V}_A$ in Model 3 was systematically smaller in datasets where model 4 did not converge than in datasets where model 4 did converge (Table S3). However, when considering only the datasets that did converge, the mean absolute error was not functionally different from the mean absolute error across all datasets, especially in comparison to the mean absolute error in model 4. The large different in accuracy between these two models under these scenarios is therefore not driven by converge problems in model 4.

Table S3: Mean absolute error across different subsets of model 3 and model 4 for scenarios 1-4.

	Model 3			Model 4
Pedigree	m4 not converged	m4 converged	all	all
ul_medium	0.0017	0.0580	0.0442	0.1265
ul_small	0.0048	0.1093	0.0887	0.2165
fl_medium	0.0011	0.0465	0.0383	0.1225
fl_small	0.0028	0.1101	0.0924	0.2137
ml_medium	0.0004	0.0514	0.0402	0.1205
ml_small	0.0032	0.1005	0.0832	0.2190
nl_medium	0.0002	0.0392	0.0309	0.1064
nl_small	0.0002	0.0912	0.0753	0.2017

S8 Full simulation results

In this section, the full results of all parameters estimated in all model are shown. Figures S18-S29 show the results from the first set of simulations aimed at assessing the bias in simple maternal effects models, with each plot showing the results from a different scenario (Table 3). Figures S30-S41 show the results from the second set of simulations aimed at comparing the performance of different models in small pedigrees, again with each plot showing a different scenario.

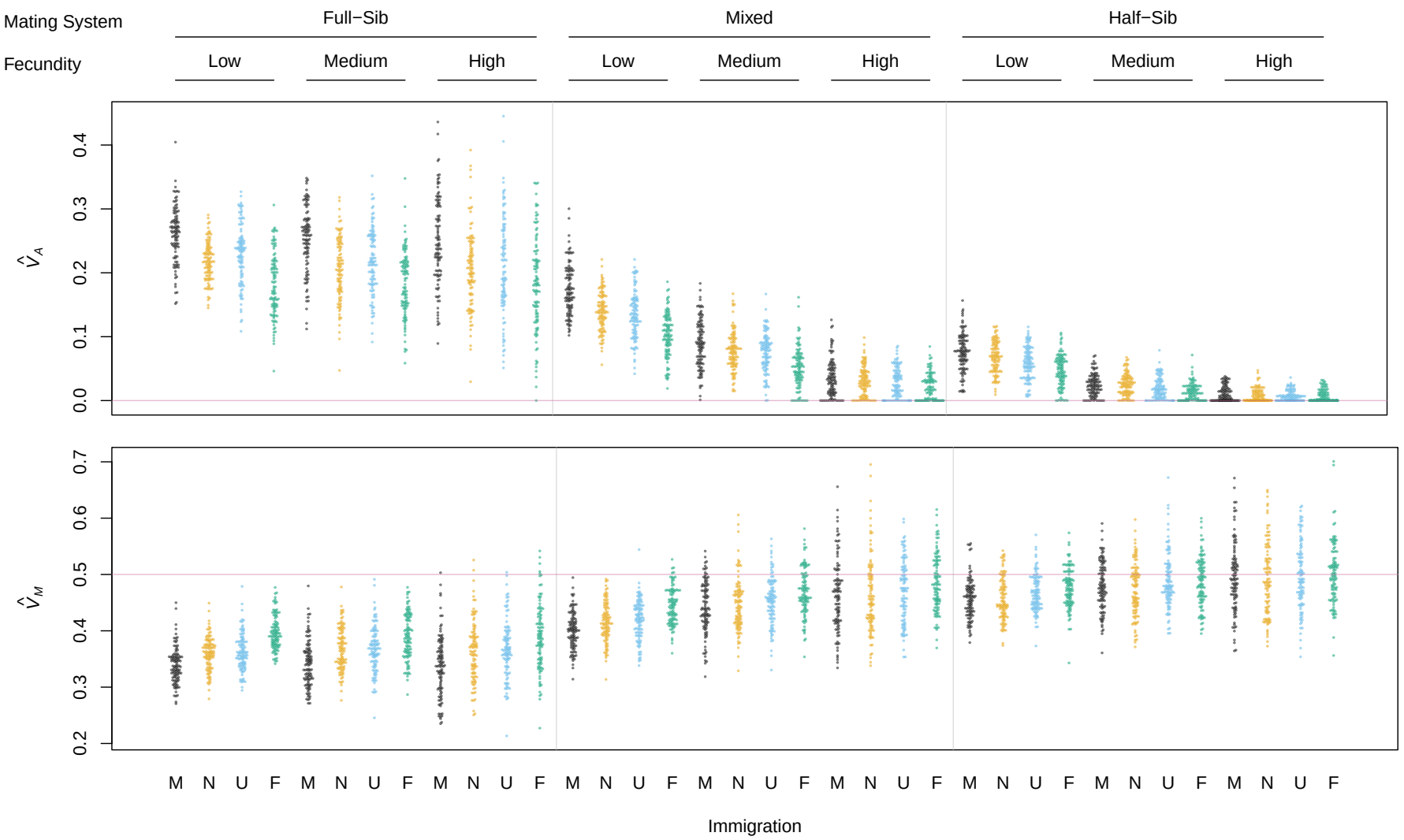


Figure S18: $\hat{V}_A$ and $\hat{V}_{Mc}$ for all simulated pedigrees from Scenario A (V_A =0, V_{Mg} =0.5, V_{Me} =0,$COV_{A,Mg}$ =0). M, N, U and F, stand for male-biased, no, unbiased and female-biased immigration respectively.

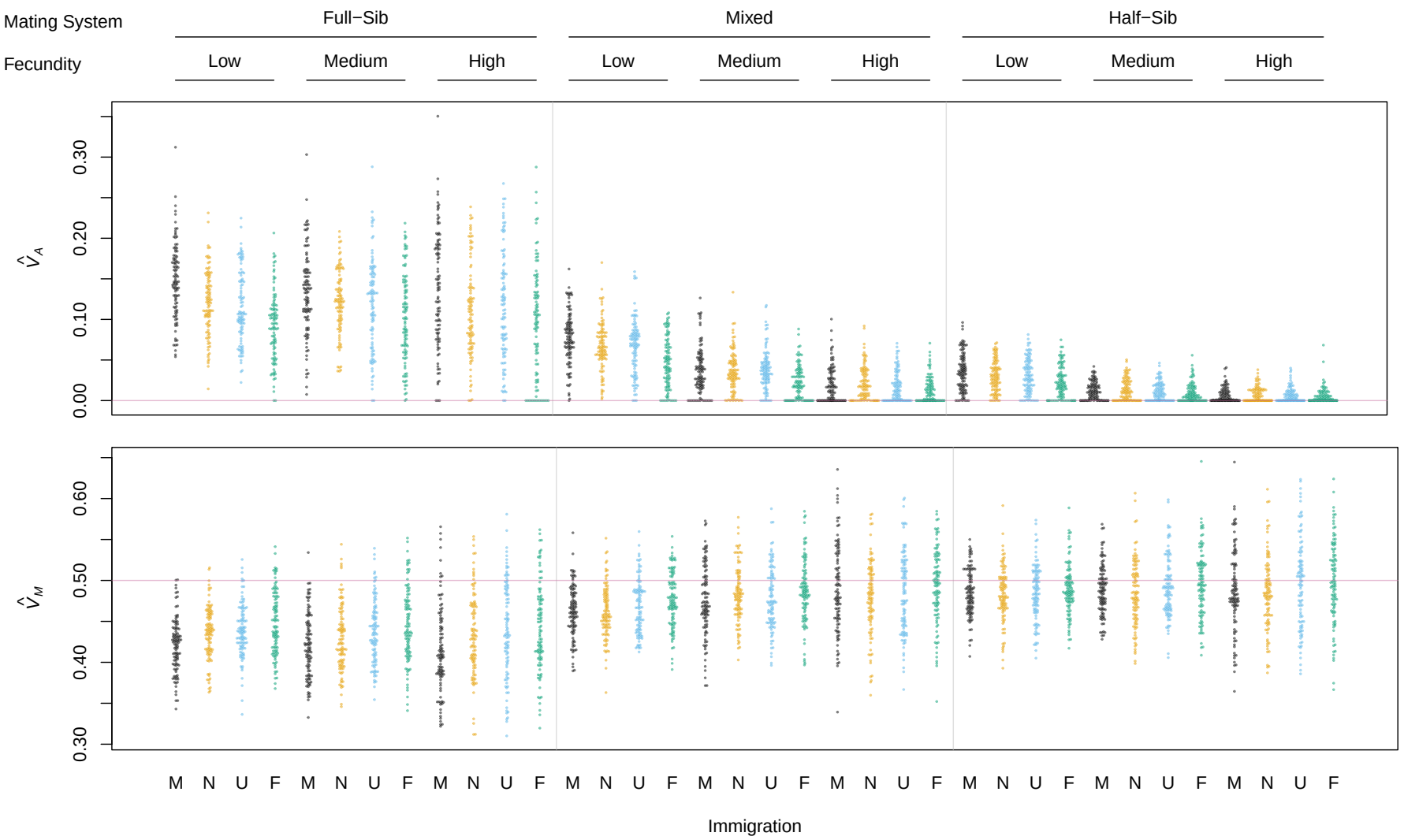


Figure S19: $\hat{V}_A$ and $\hat{V}_{Mc}$ for all simulated pedigrees from Scenario B (V_A =0, V_{Mg} =0.25, V_{Me} =0.25,$COV_{A,Mg}$ =0). M, N, U and F, stand for male-biased, no, unbiased and female-biased immigration respectively.

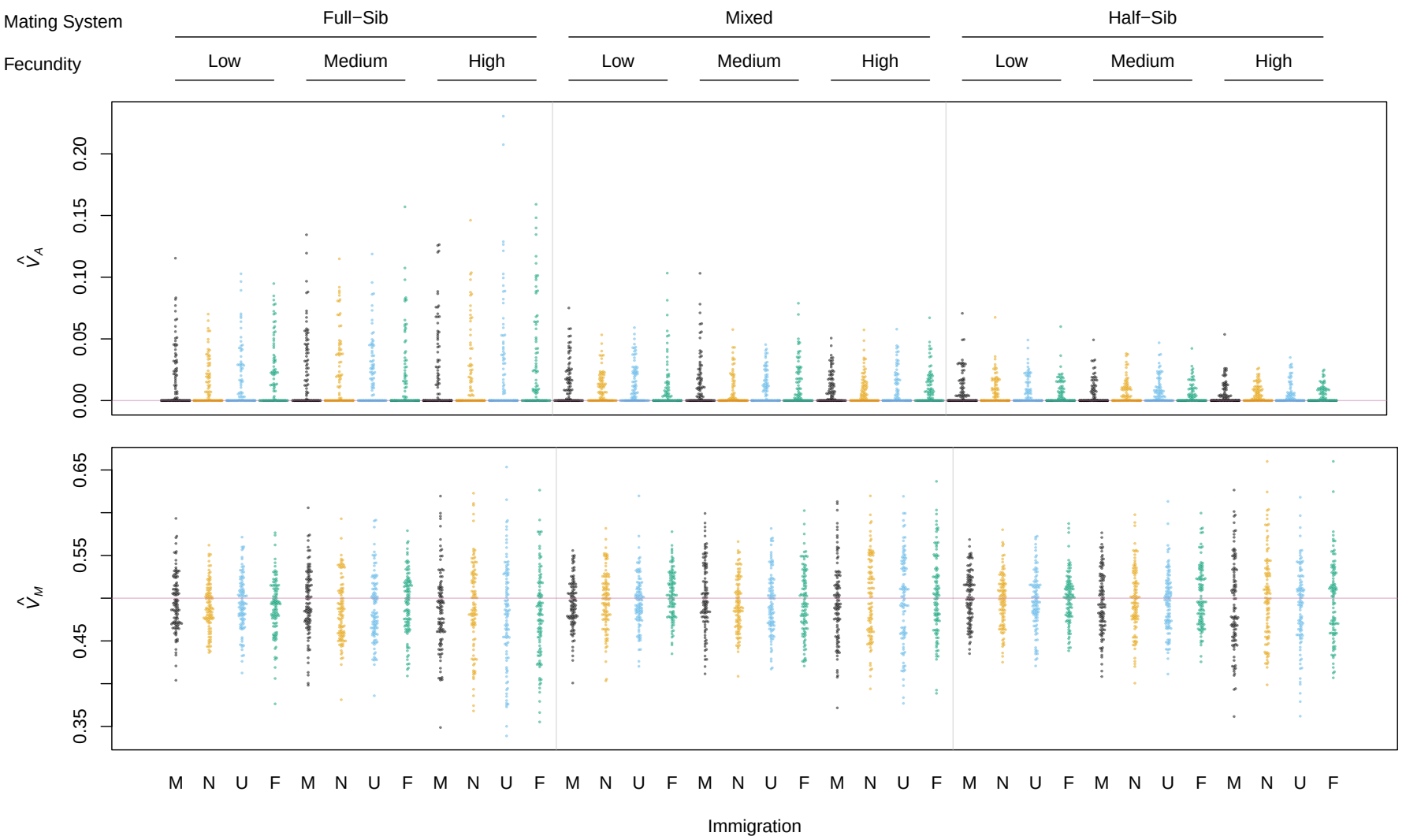


Figure S20: $\hat{V}_A$ and $\hat{V}_{Mc}$ for all simulated pedigrees from Scenario C (V_A =0, V_{Mg} =0, V_{Me} =0.5,$COV_{A,Mg}$ =0). M, N, U and F, stand for male-biased, no, unbiased and female-biased immigration respectively.

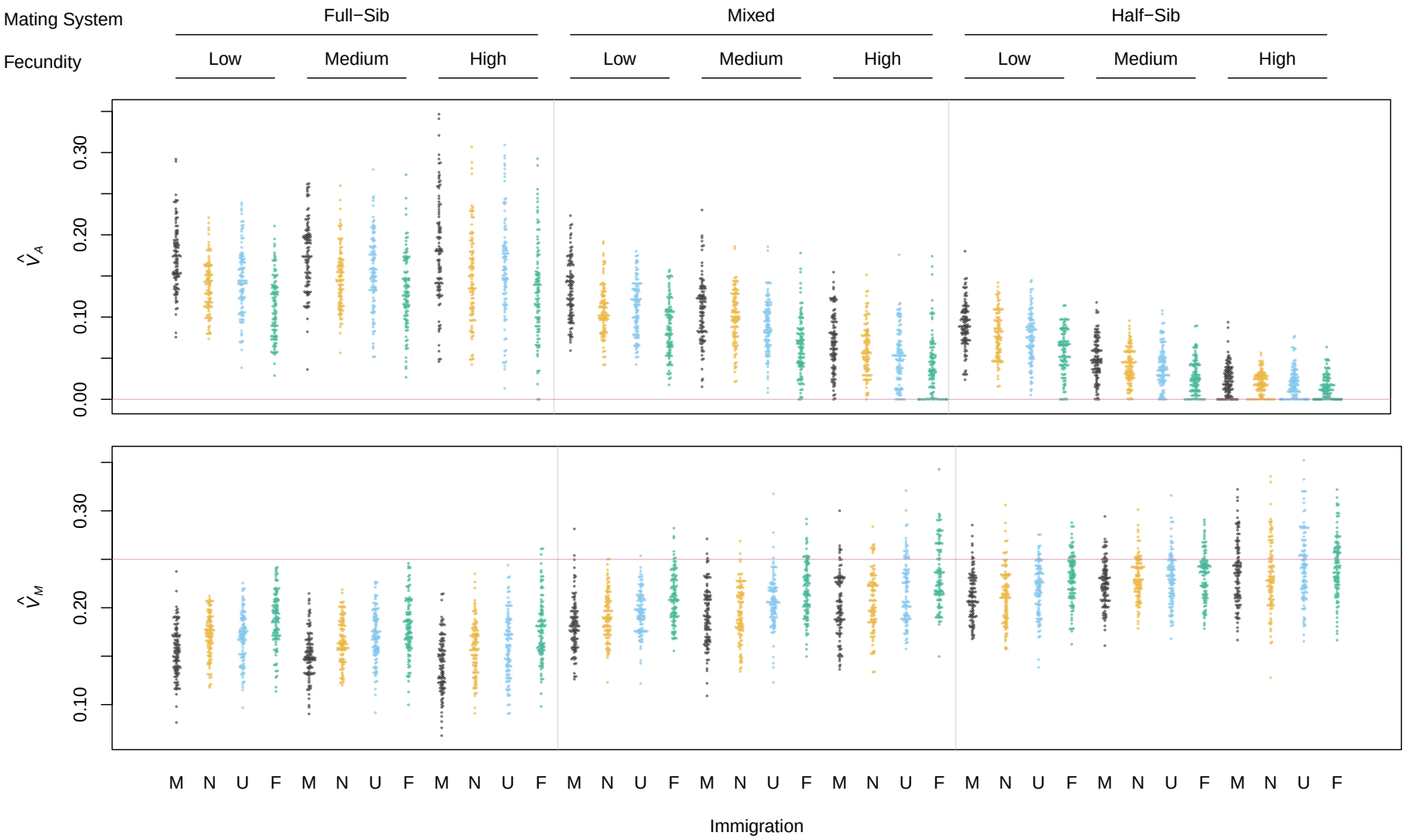


Figure S21: $\hat{V}_A$ and $\hat{V}_{Mc}$ for all simulated pedigrees from Scenario D (V_A =0, V_{Mg} =0.25, V_{Me} =0,$COV_{A,Mg}$ =0). M, N, U and F, stand for male-biased, no, unbiased and female-biased immigration respectively.

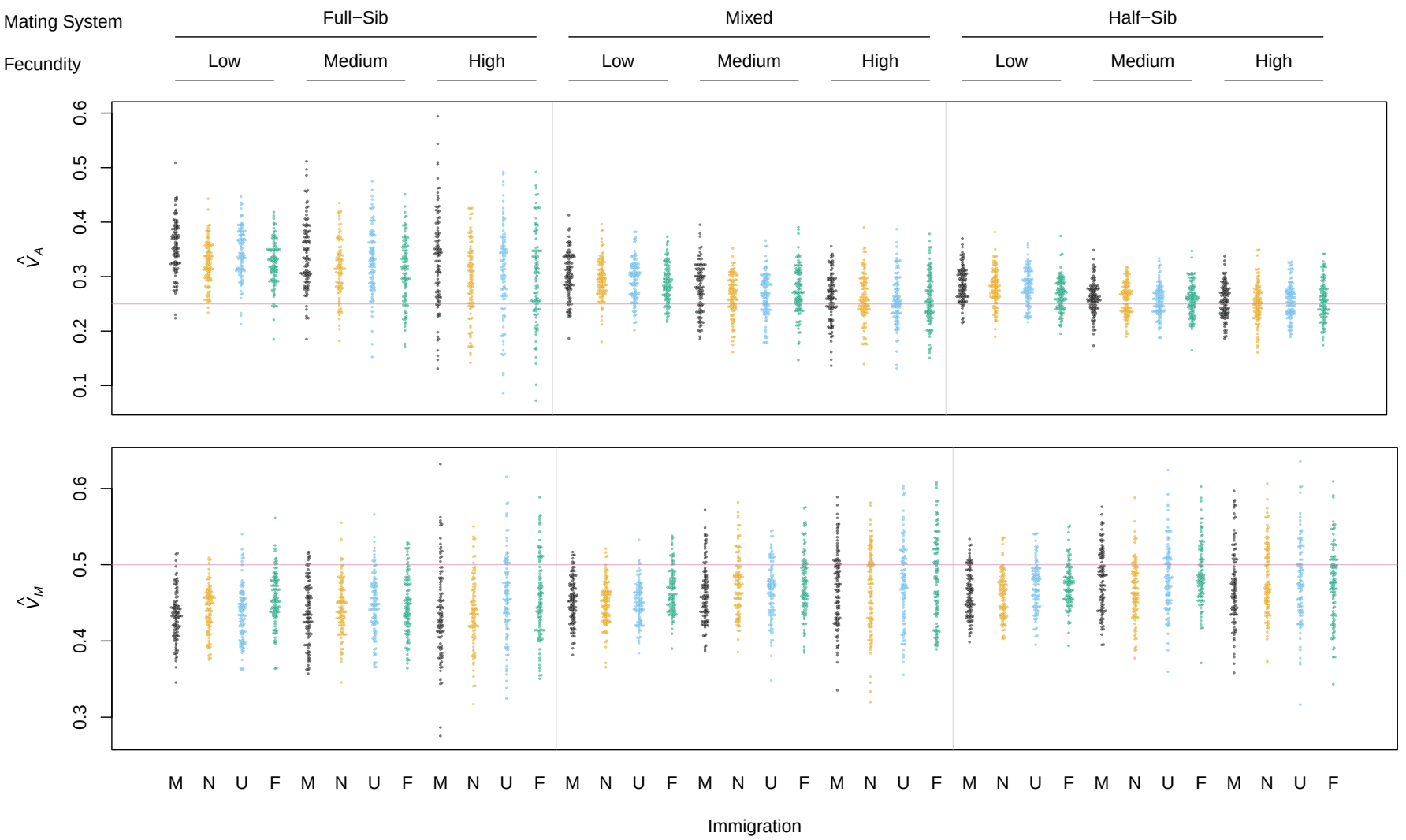


Figure S22: $\hat{V}_A$ and $\hat{V}_{Mc}$ for all simulated pedigrees from Scenario E (V_A =0.25, V_{Mg} =0.25, V_{Me} =0.25,$COV_{A,Mg}$ =0). M, N, U and F, stand for male-biased, no, unbiased and female-biased immigration respectively.

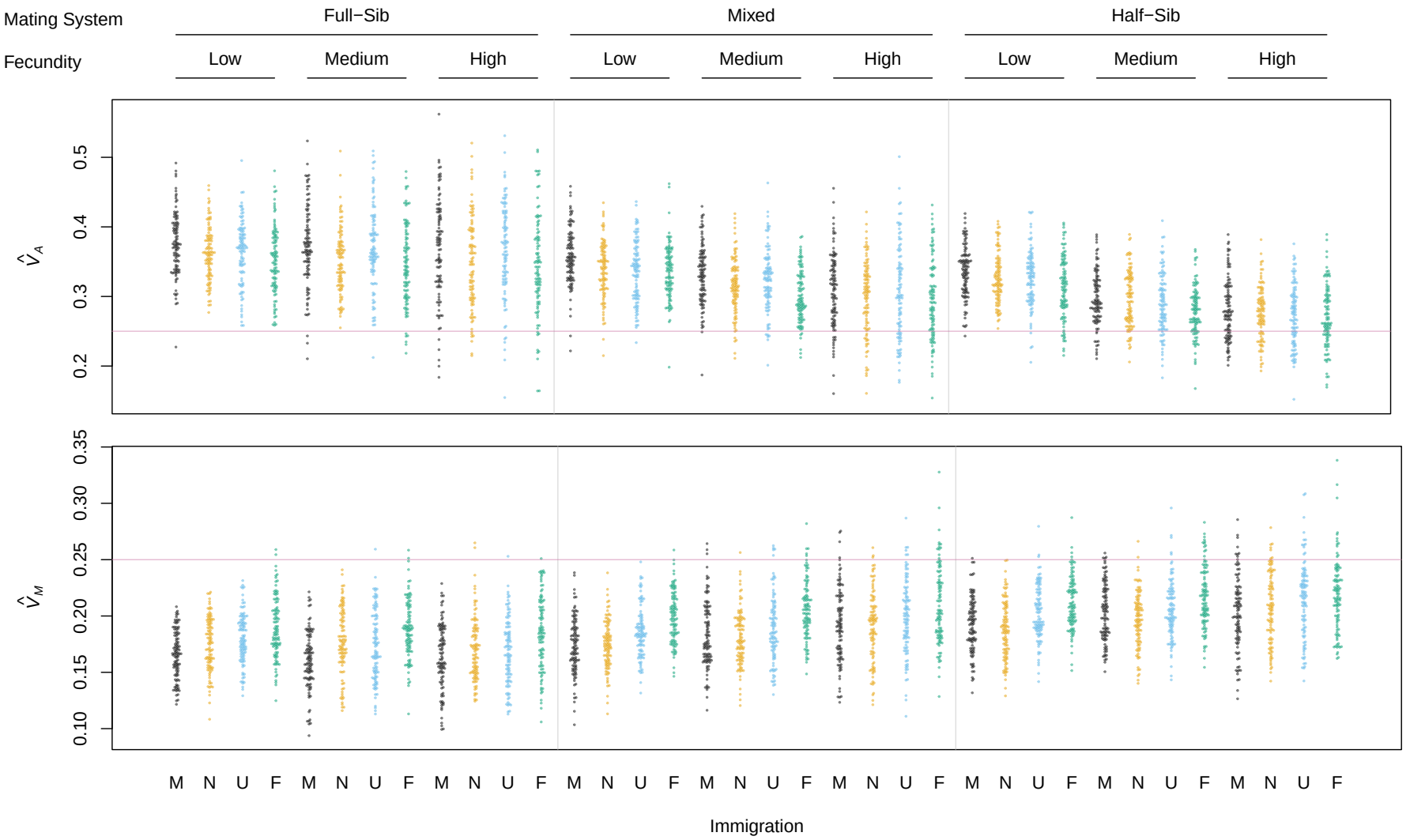


Figure S23: $\hat{V}_A$ and $\hat{V}_{Mc}$ for all simulated pedigrees from Scenario F (V_A =0.25, V_{Mg} =0.25, V_{Me} =0,$COV_{A,Mg}$ =0). M, N, U and F, stand for male-biased, no, unbiased and female-biased immigration respectively.

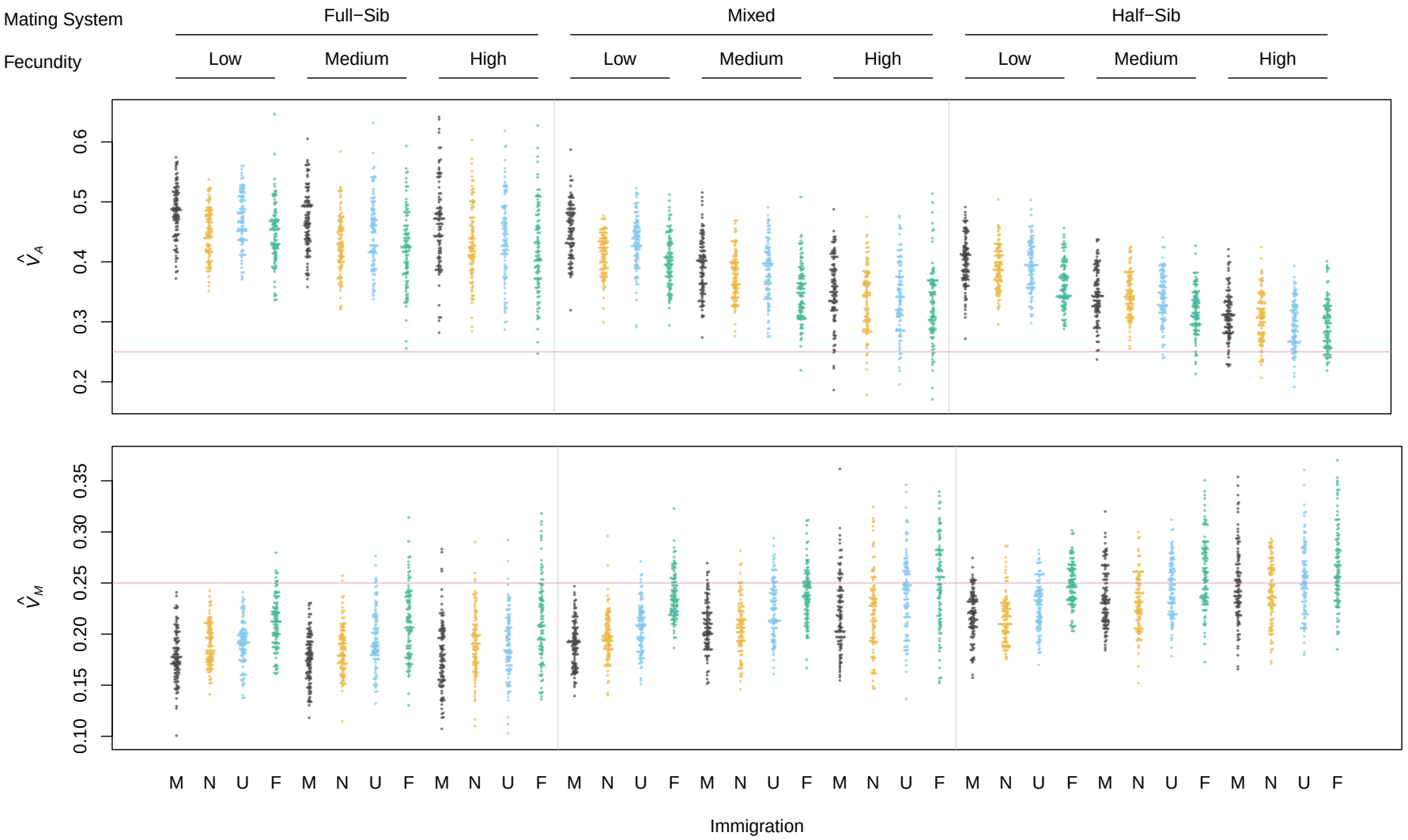


Figure S24: $\hat{V}_A$ and $\hat{V}_{Mc}$ for all simulated pedigrees from Scenario G (V_A =0.25, V_{Mg} =0.25, V_{Me} =0,$COV_{A,Mg}$ =0.075). M, N, U and F, stand for male-biased, no, unbiased and female-biased immigration respectively.

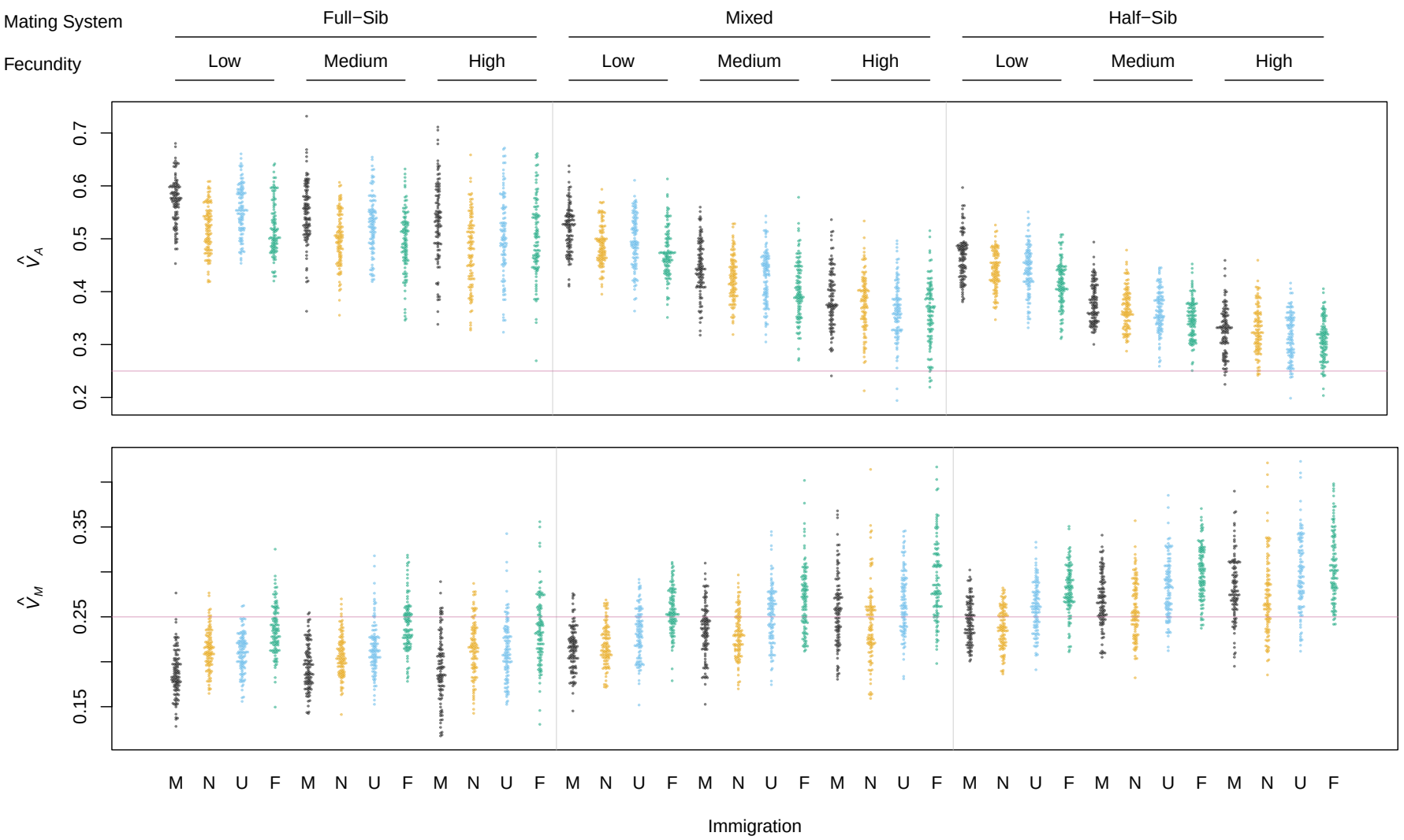


Figure S25: $\hat{V}_A$ and $\hat{V}_{Mc}$ for all simulated pedigrees from Scenario H (V_A =0.25, V_{Mg} =0.25, V_{Me} =0,$COV_{A,Mg}$ =0.15). M, N, U and F, stand for male-biased, no, unbiased and female-biased immigration respectively.

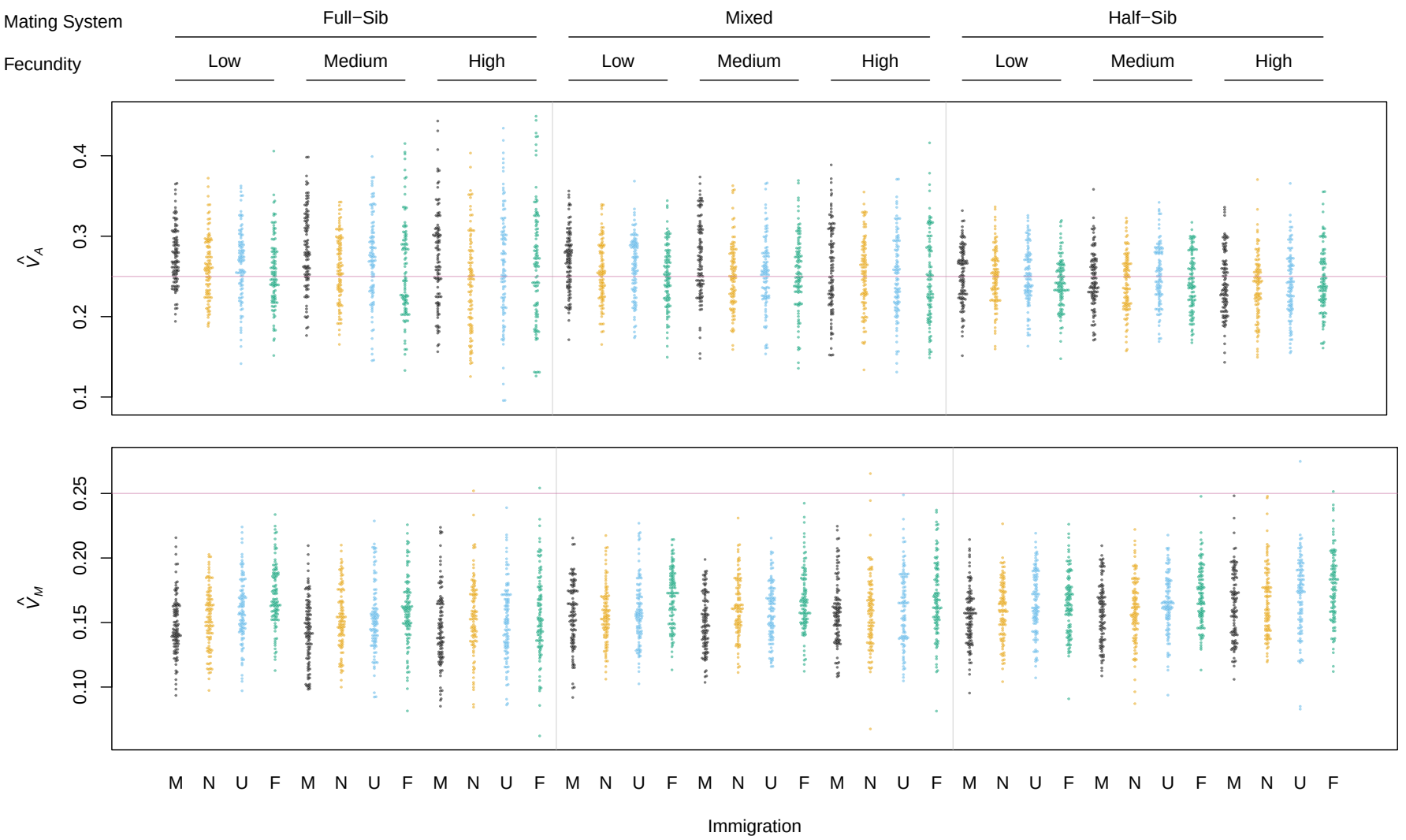


Figure S26: $\hat{V}_A$ and $\hat{V}_{Mc}$ for all simulated pedigrees from Scenario I (V_A =0.25, V_{Mg} =0.25, V_{Me} =0,$COV_{A,Mg}$ =-0.075). M, N, U and F, stand for male-biased, no, unbiased and female-biased immigration respectively.

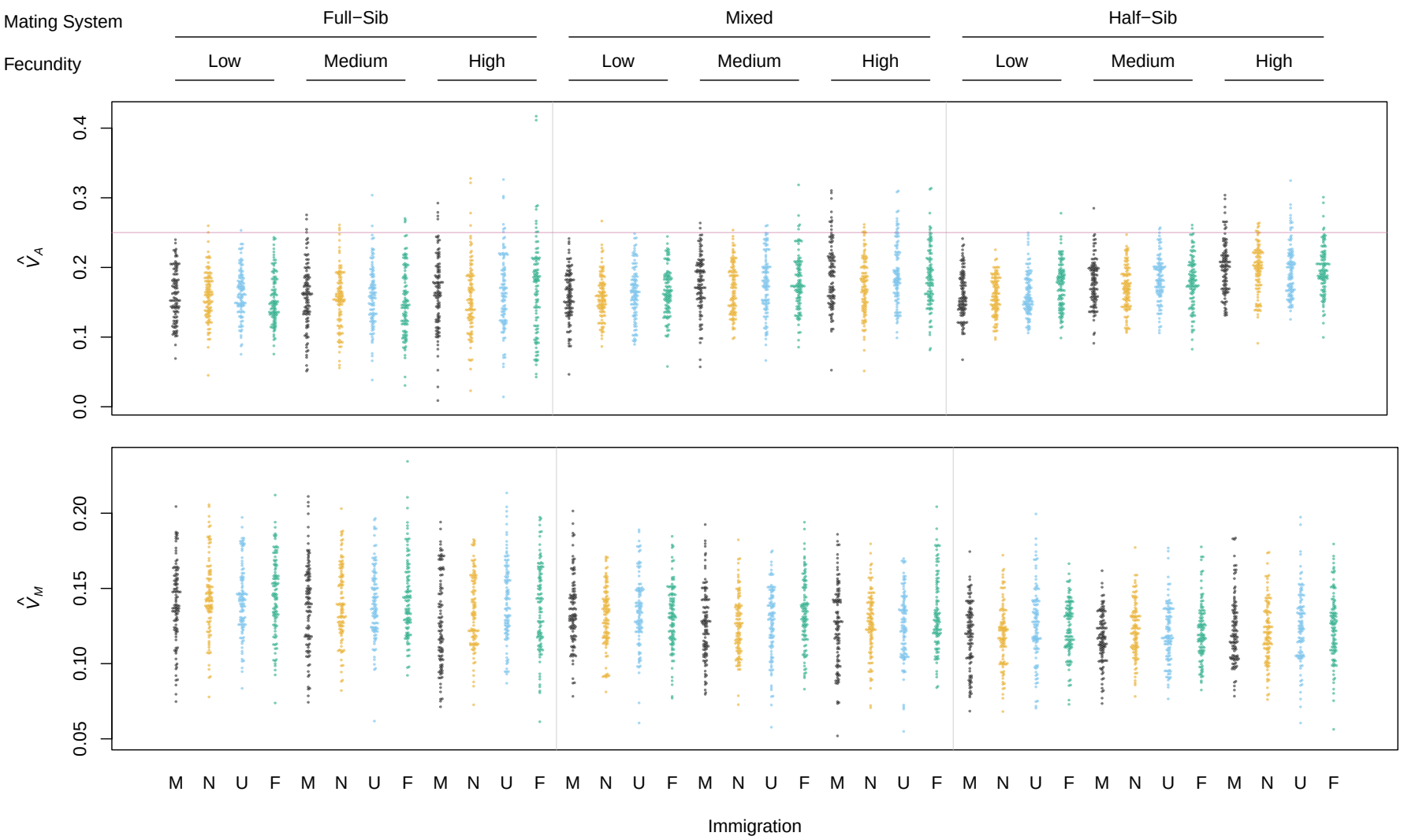


Figure S27: $\hat{V}_A$ and $\hat{V}_{Mc}$ for all simulated pedigrees from Scenario J (V_A =0.25, V_{Mg} =0.25, V_{Me} =0,$COV_{A,Mg}$ =-0.15). M, N, U and F, stand for male-biased, no, unbiased and female-biased immigration respectively.

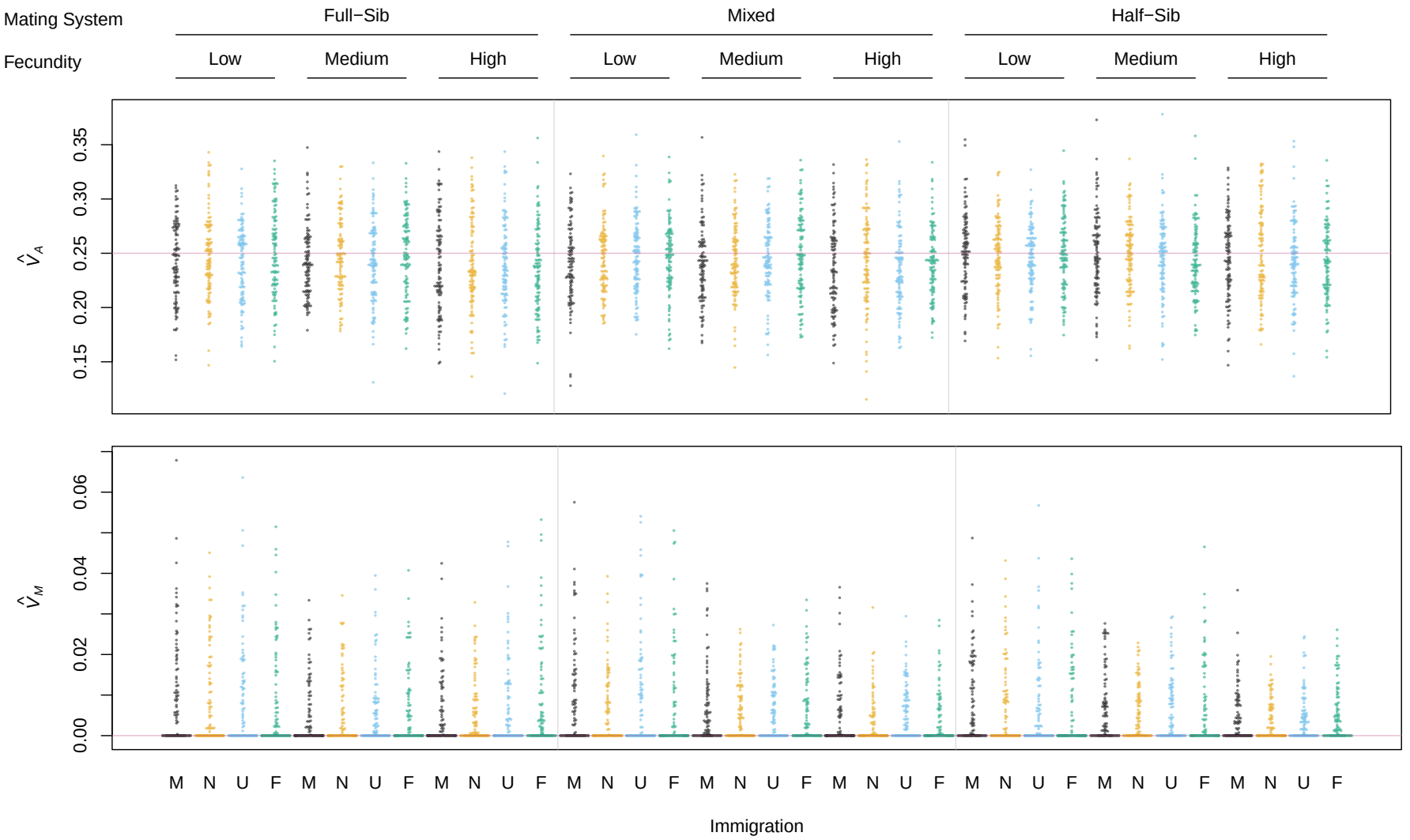


Figure S28: $\hat{V}_A$ and $\hat{V}_{Mc}$ for all simulated pedigrees from Scenario K (V_A =0.25, V_{Mg} =0, V_{Me} =0,$COV_{A,Mg}$ =0). M, N, U and F, stand for male-biased, no, unbiased and female-biased immigration respectively.

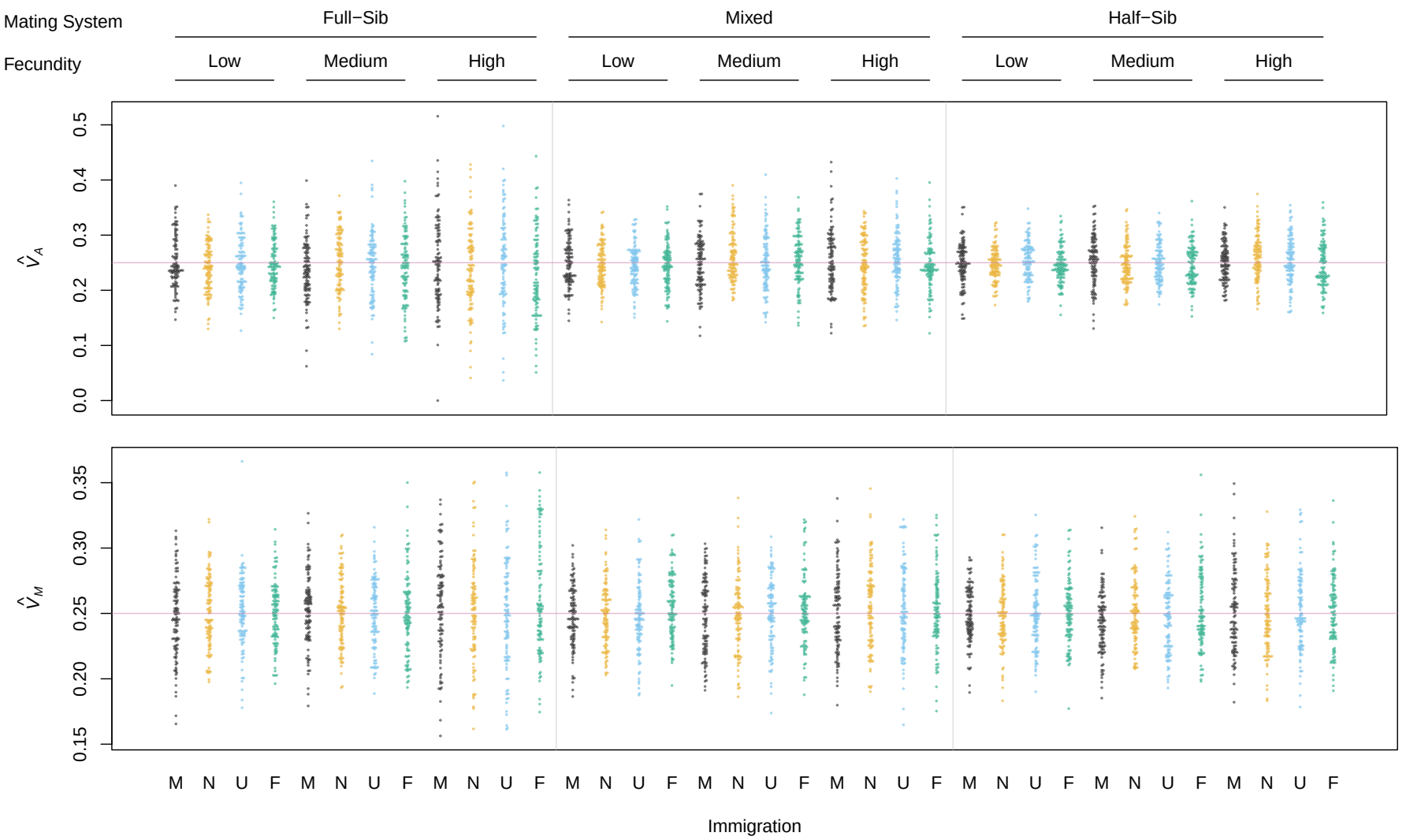


Figure S29: $\hat{V}_A$ and $\hat{V}_{Mc}$ for all simulated pedigrees from Scenario L (V_A =0.25, V_{Mg} =0, V_{Me} =0.25,$COV_{A,Mg}$ =0). M, N, U and F, stand for male-biased, no, unbiased and female-biased immigration respectively.

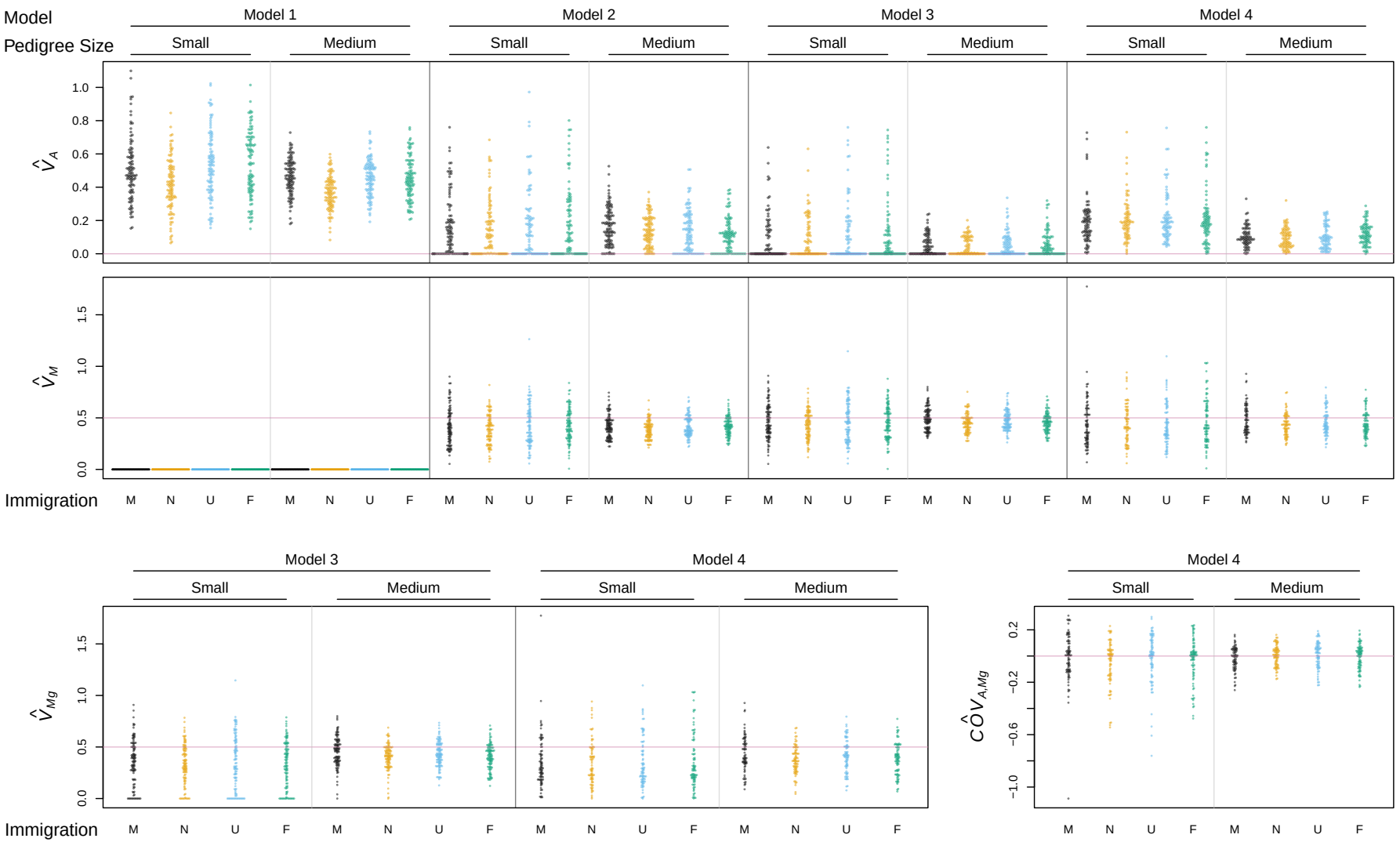


Figure S30: $\hat{V}_A$, $\hat{V}_M$, $\hat{V}_{Mg}$ and $C\hat{O}V_{A,Mg}$ from small and medium sized pedigree simulations across all simulated pedigrees from Scenario A (V_A =0, V_{Mg} =0.5, V_{Me} =0,$COV_{A,Mg}$ =0). M, N, U and F, stand for male-biased, no, unbiased and female-biased immigration respectively.

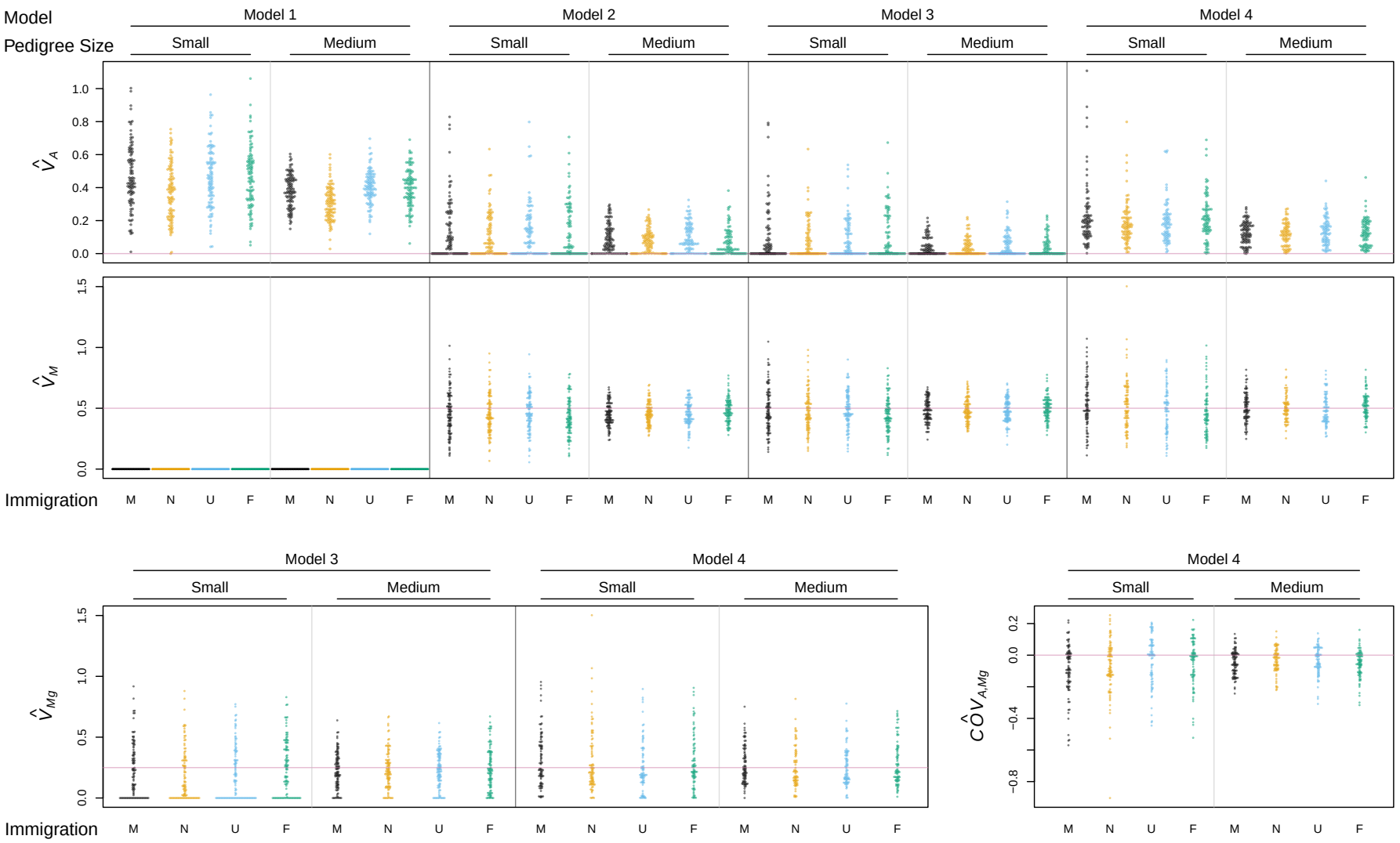


Figure S31: $\hat{V}_A$, $\hat{V}_M$, $\hat{V}_{Mg}$ and $C\hat{O}V_{A,Mg}$ from small and medium sized pedigree simulations across all simulated pedigrees from Scenario B (V_A =0, V_{Mg} =0.25, V_{Me} =0.25,$COV_{A,Mg}$ =0). M, N, U and F, stand for male-biased, no, unbiased and female-biased immigration respectively.

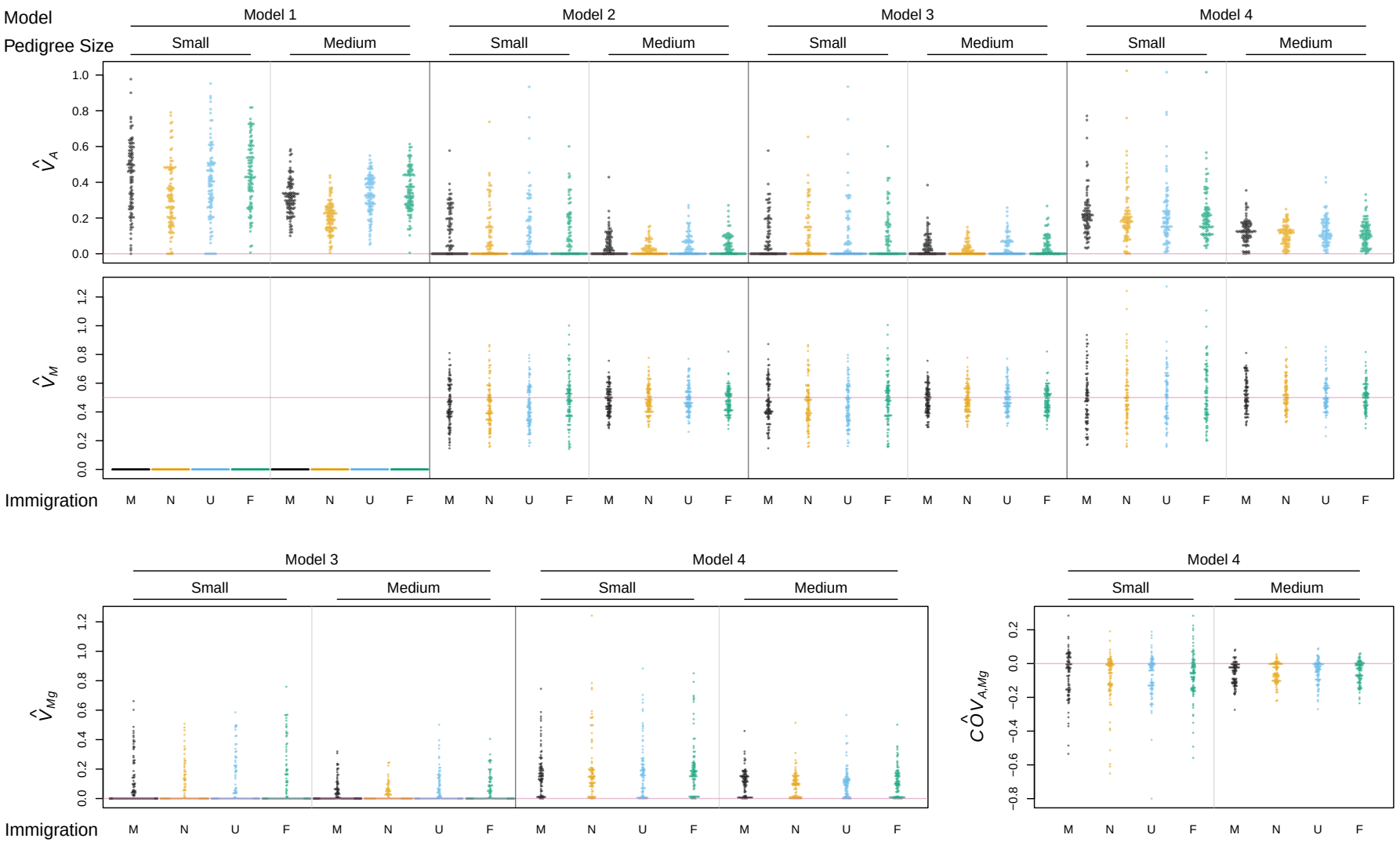


Figure S32: $\hat{V}_A$, $\hat{V}_M$, $\hat{V}_{Mg}$ and $C\hat{O}V_{A,Mg}$ from small and medium sized pedigree simulations across all simulated pedigrees Scenario C (V_A =0, V_{Mg} =0, V_{Me} =0.5,$COV_{A,Mg}$ =0). M, N, U and F, stand for male-biased, no, unbiased and female-biased immigration respectively.

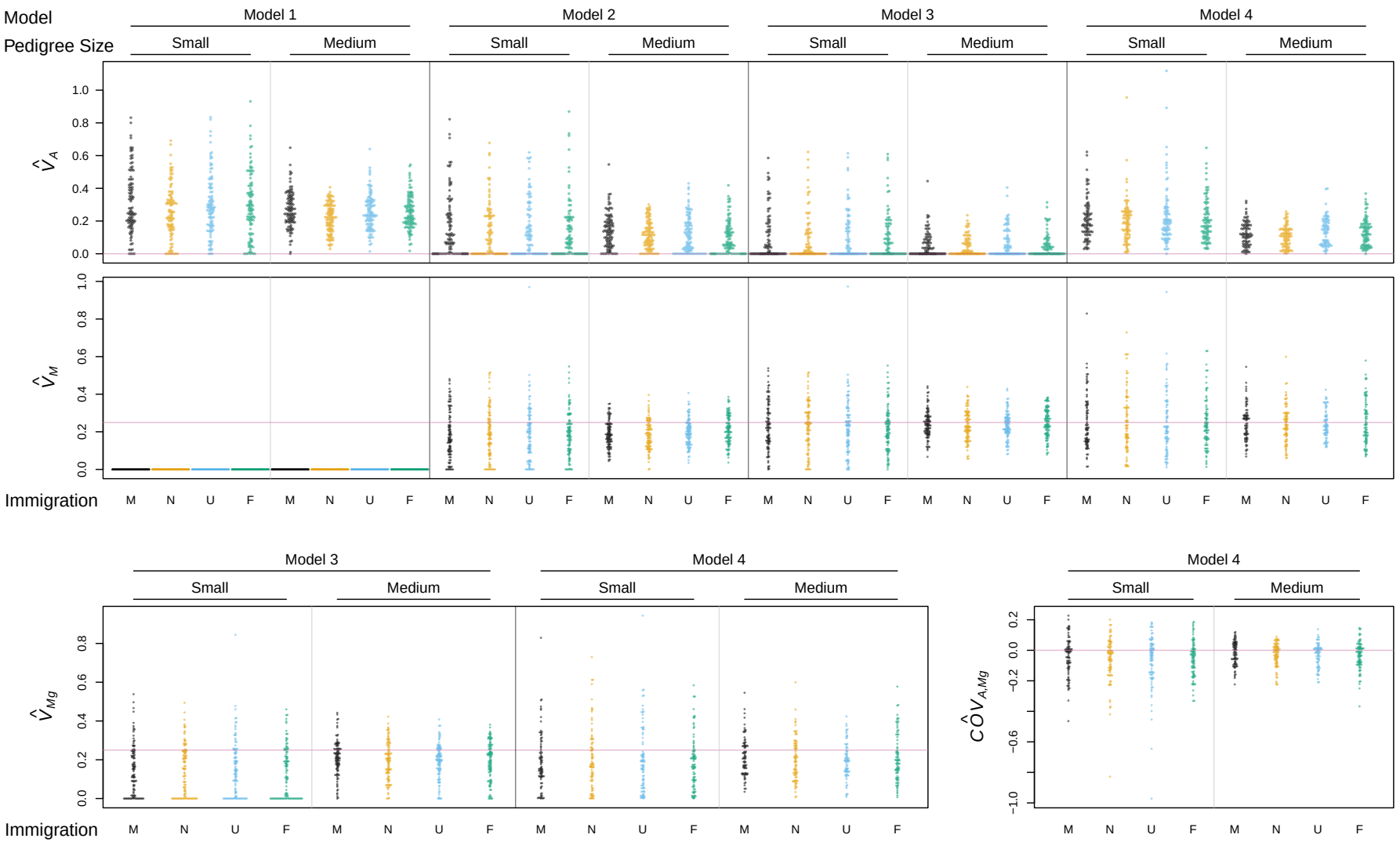


Figure S33: $\hat{V}_A$, $\hat{V}_M$, $\hat{V}_{Mg}$ and $C\hat{O}V_{A,Mg}$ from small and medium sized pedigree simulations across all simulated pedigrees from Scenario D (V_A =0, V_{Mg} =0.25, V_{Me} =0,$COV_{A,Mg}$ =0). M, N, U and F, stand for male-biased, no, unbiased and female-biased immigration respectively.

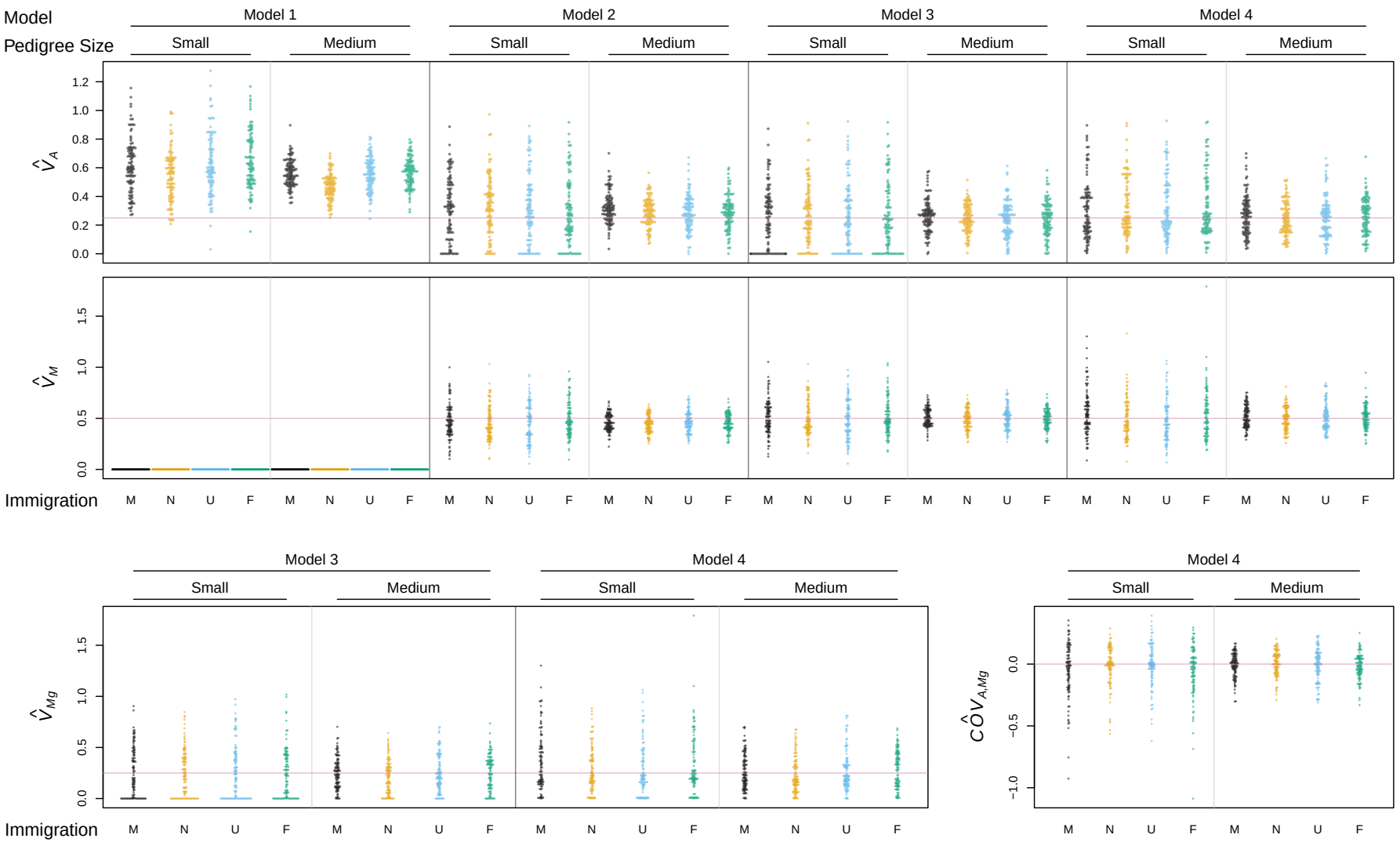


Figure S34: $\hat{V}_A$, $\hat{V}_M$, $\hat{V}_{Mg}$ and $C\hat{O}V_{A,Mg}$ from small and medium sized pedigree simulations across all simulated pedigrees from Scenario E (V_A =0.25, V_{Mg} =0.25, V_{Me} =0.25,$COV_{A,Mg}$ =0). M, N, U and F, stand for male-biased, no, unbiased and female-biased immigration respectively.

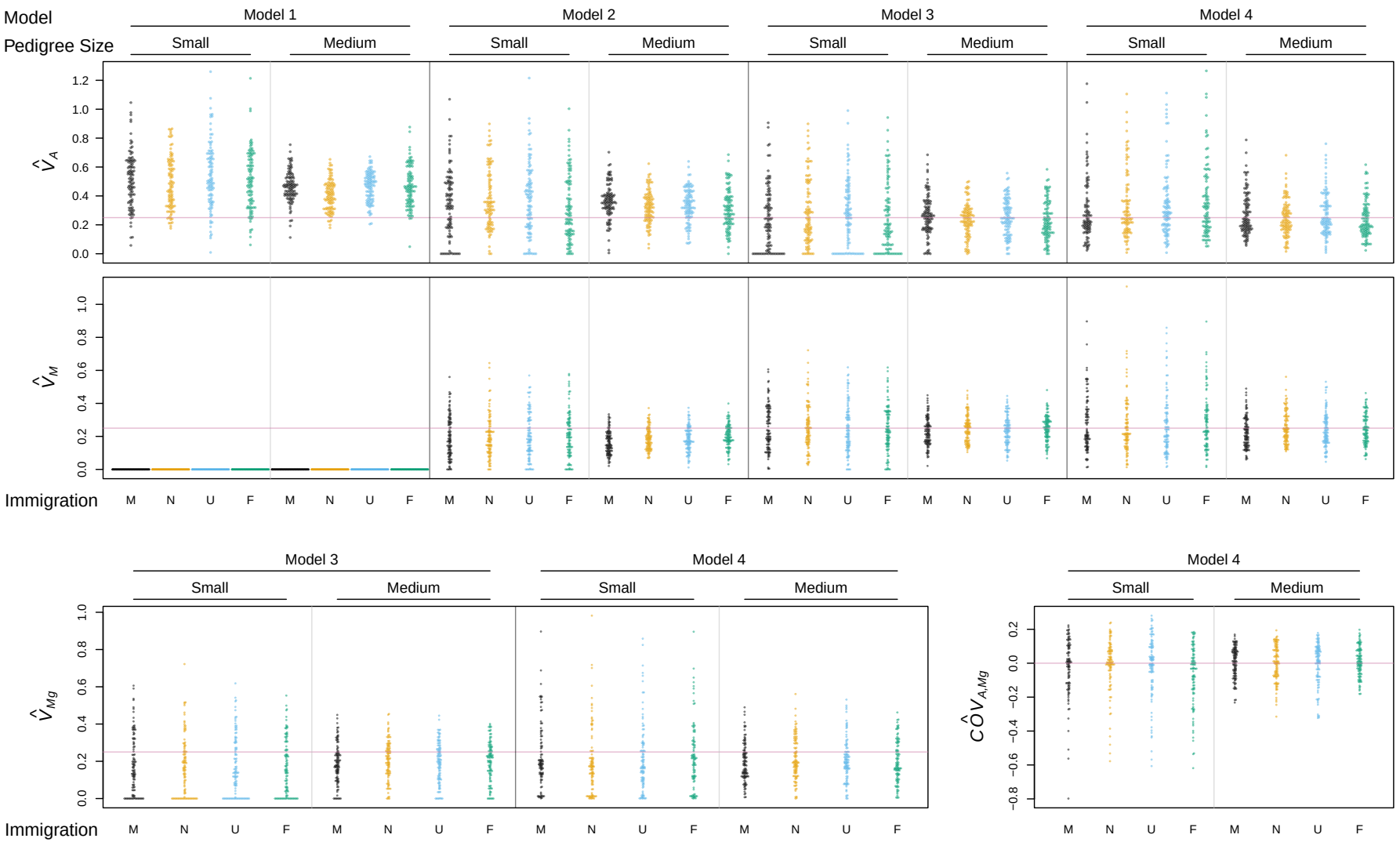


Figure S35: $\hat{V}_A$, $\hat{V}_M$, $\hat{V}_{Mg}$ and $C\hat{O}V_{A,Mg}$ from small and medium sized pedigree simulations across all simulated pedigrees from Scenario F (V_A =0.25, V_{Mg} =0.25, V_{Me} =0,$COV_{A,Mg}$ =0). M, N, U and F, stand for male-biased, no, unbiased and female-biased immigration respectively.

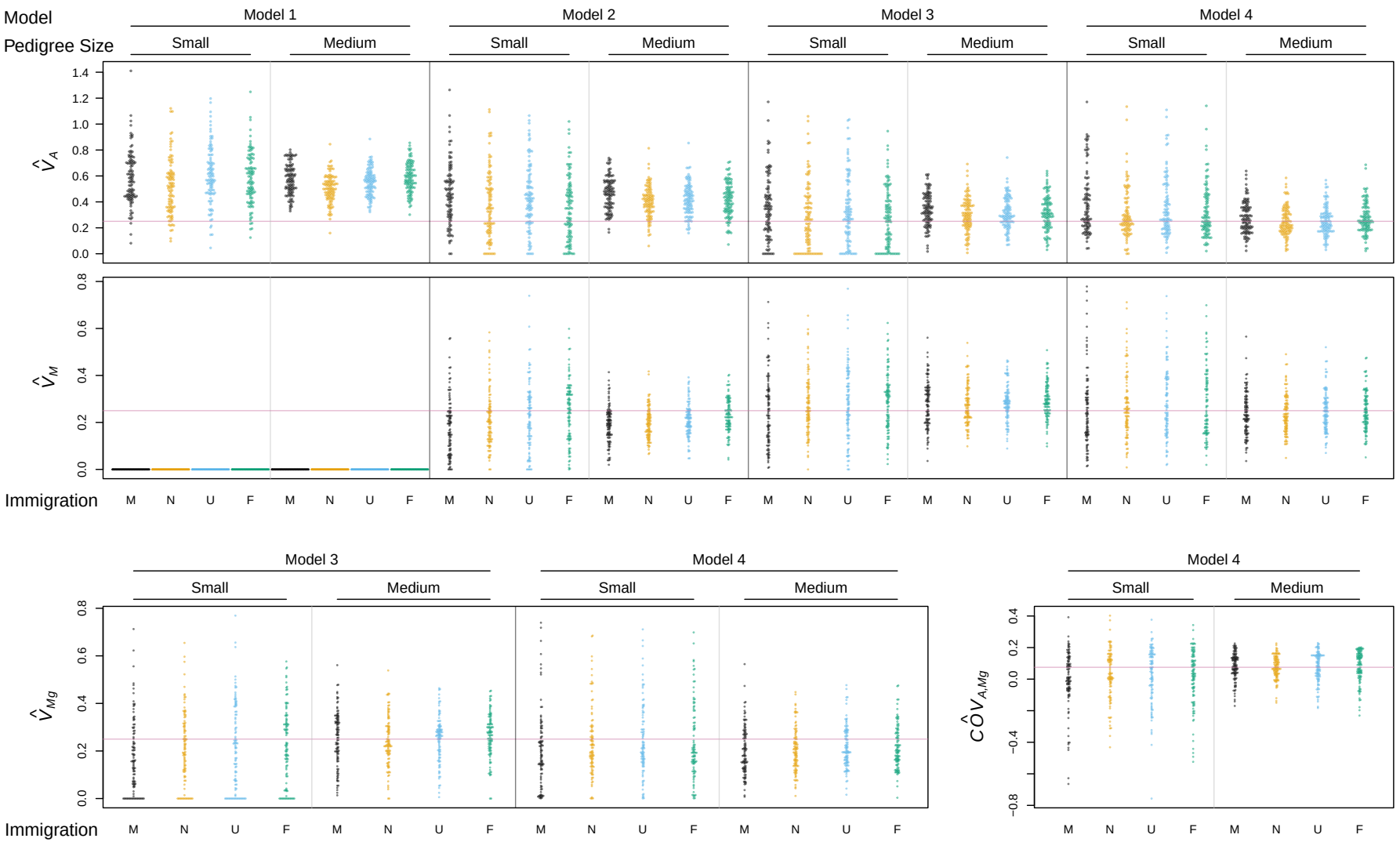


Figure S36: $\hat{V}_A$, $\hat{V}_M$, $\hat{V}_{Mg}$ and $\hat{COV}_{A,Mg}$ from small and medium sized pedigree simulations across all simulated pedigrees from Scenario G (V_A =0.25, V_{Mg} =0.25, V_{Me} =0,$COV_{A,Mg}$ =0.075). M, N, U and F, stand for male-biased, no, unbiased and female-biased immigration respectively.

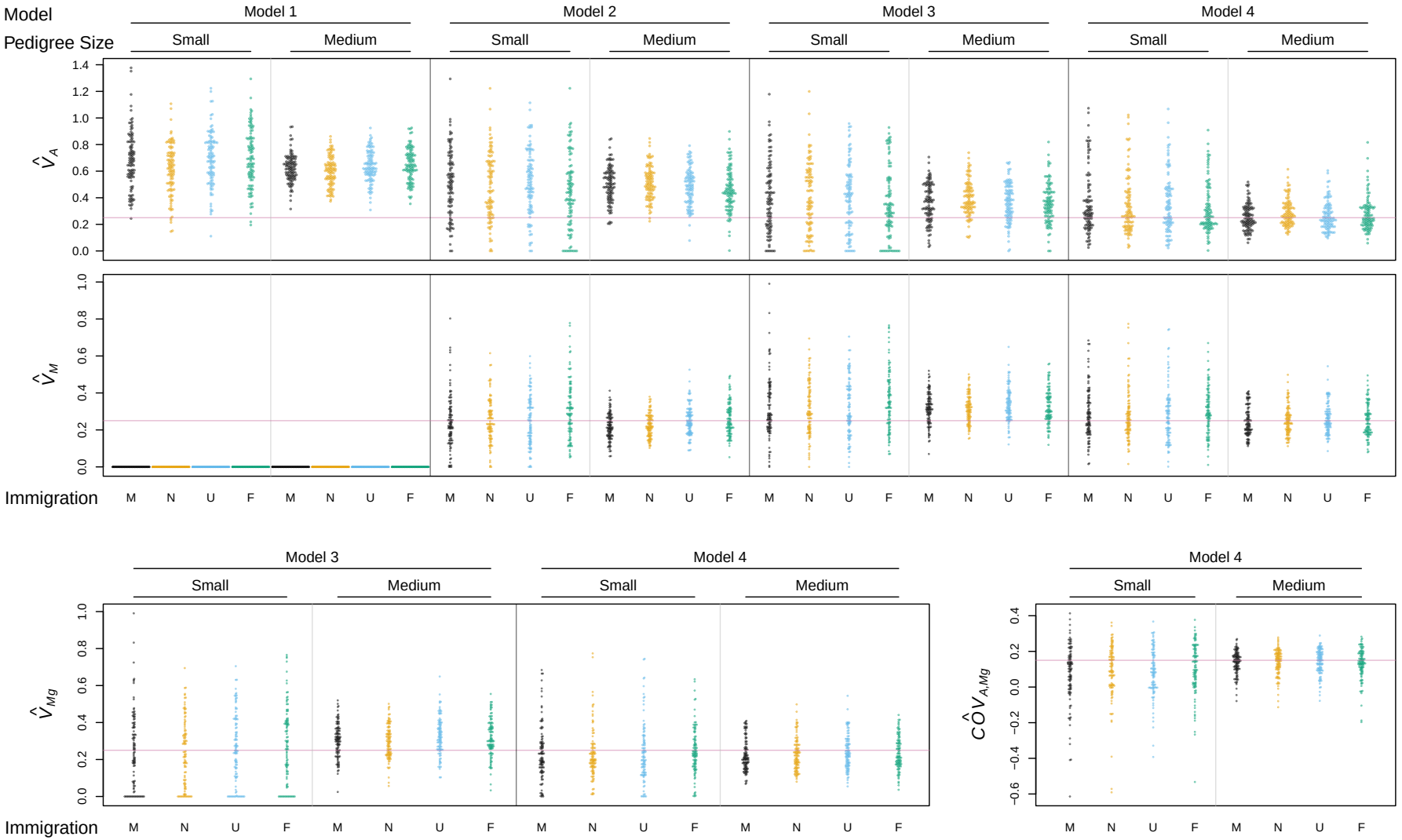


Figure S37: $\hat{V}_A$, $\hat{V}_M$, $\hat{V}_{Mg}$ and $\hat{COV}_{A,Mg}$ from small and medium sized pedigree simulations across all simulated pedigrees from Scenario H (V_A =0.25, V_{Mg} =0.25, V_{Me} =0,$COV_{A,Mg}$ =0.15). M, N, U and F, stand for male-biased, no, unbiased and female-biased immigration respectively.

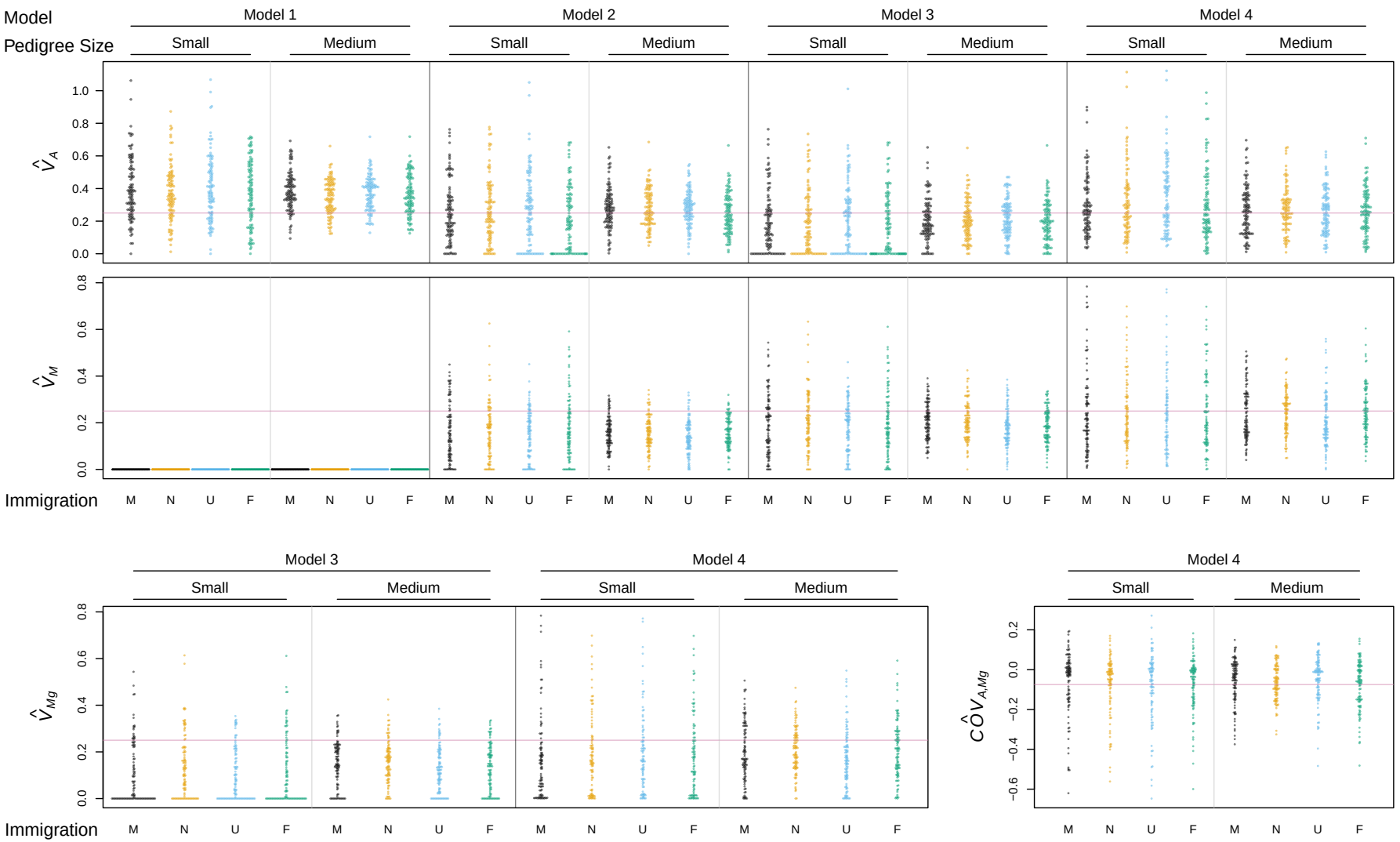


Figure S38: $\hat{V}_A$, $\hat{V}_M$, $\hat{V}_{Mg}$ and $C\hat{O}V_{A,Mg}$ from small and medium sized pedigree simulations across all simulated pedigrees from Scenario I (V_A =0.25, V_{Mg} =0.25, V_{Me} =0,$COV_{A,Mg}$ =-0.075). M, N, U and F, stand for male-biased, no, unbiased and female-biased immigration respectively.

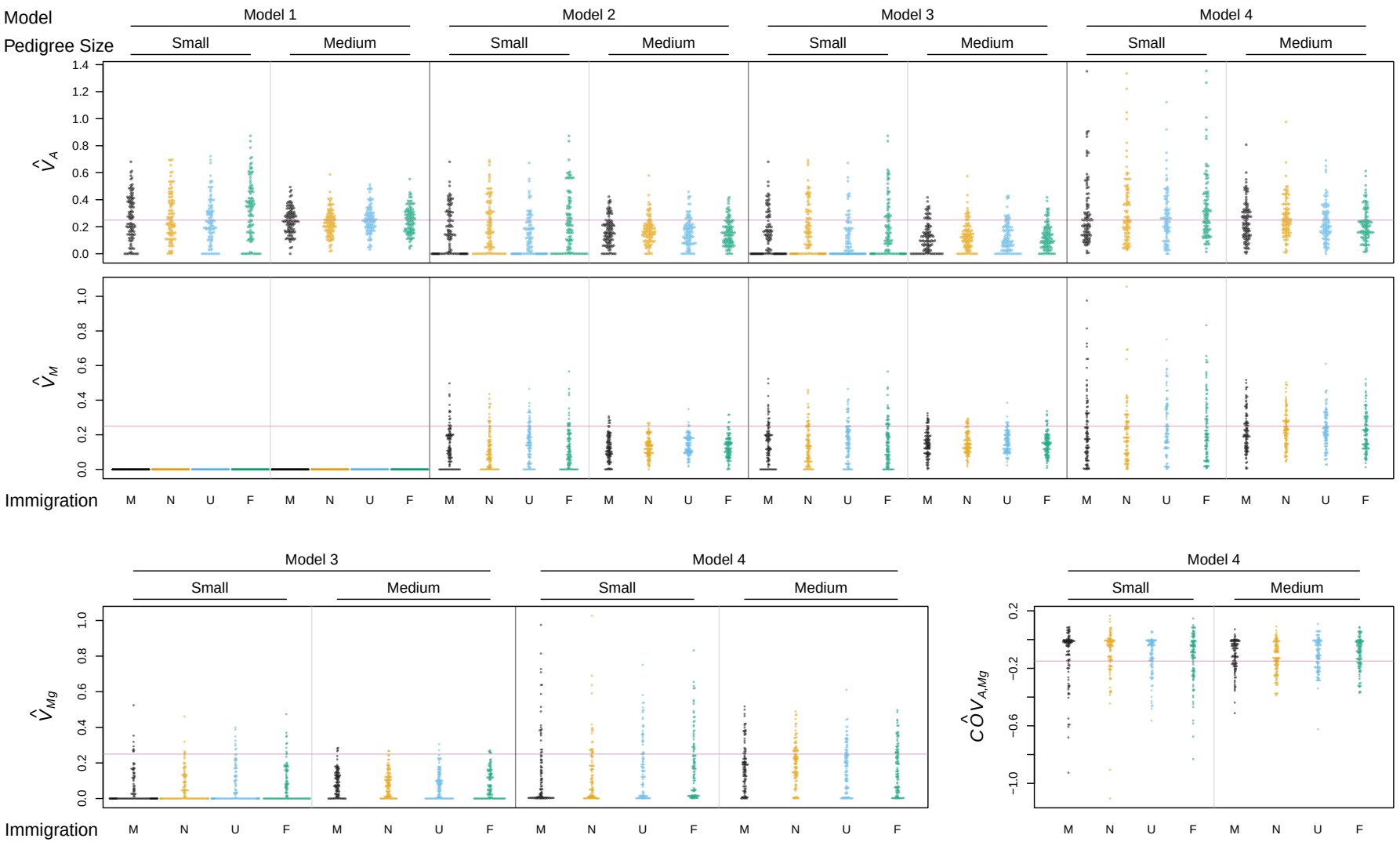


Figure S39: $\hat{V}_A$, $\hat{V}_M$, $\hat{V}_{Mg}$ and $\hat{COV}_{A,Mg}$ from small and medium sized pedigree simulations across all simulated pedigrees from Scenario J (V_A =0.25, V_{Mg} =0.25, V_{Me} =0,$COV_{A,Mg}$ =-0.15). M, N, U and F, stand for male-biased, no, unbiased and female-biased immigration respectively.

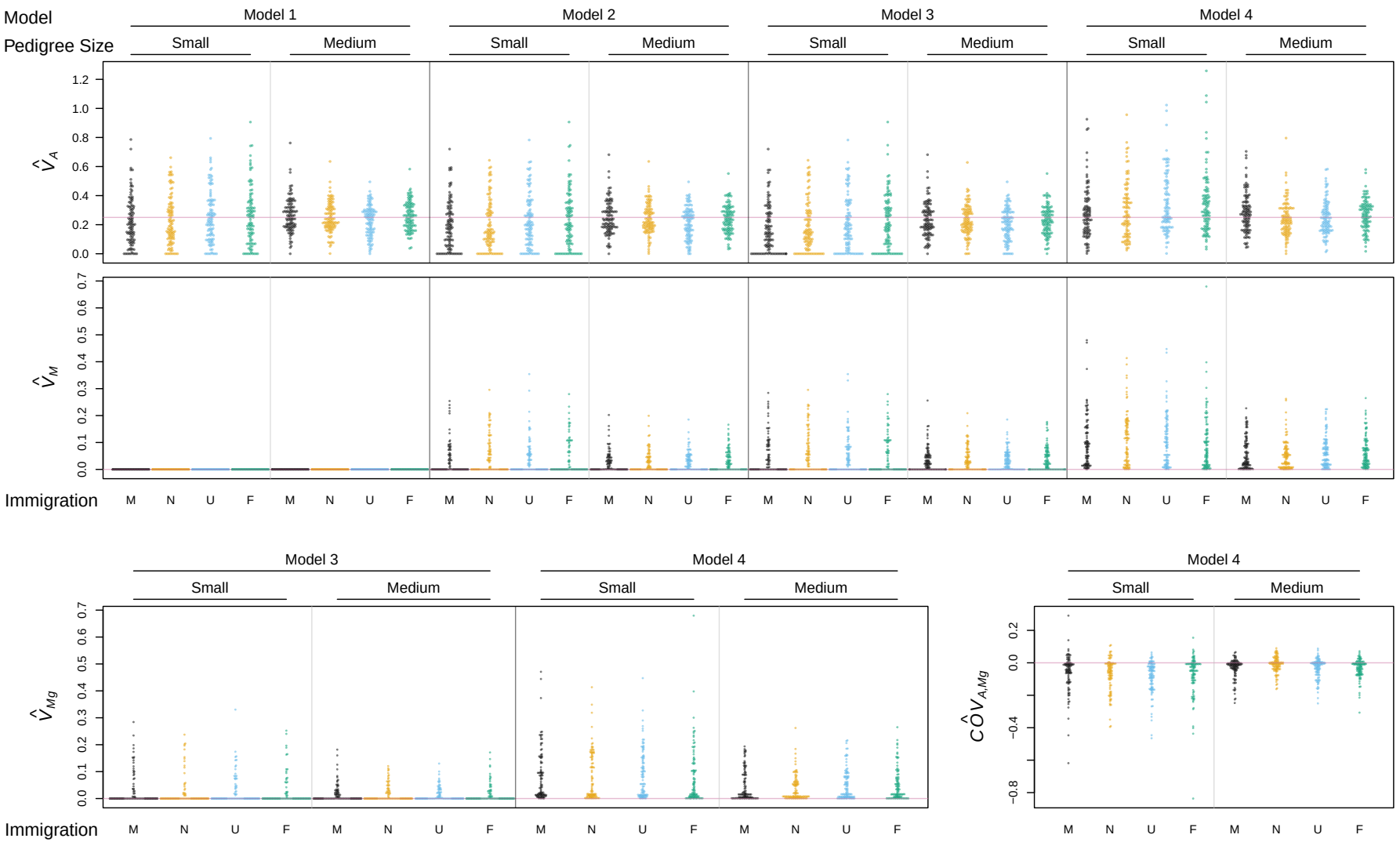


Figure S40: $\hat{V}_A$, $\hat{V}_M$, $\hat{V}_{Mg}$ and $C\hat{O}V_{A,Mg}$ from small and medium sized pedigree simulations across all simulated pedigrees from Scenario K (V_A =0.25, V_{Mg} =0, V_{Me} =0,$COV_{A,Mg}$ =0). M, N, U and F, stand for male-biased, no, unbiased and female-biased immigration respectively.

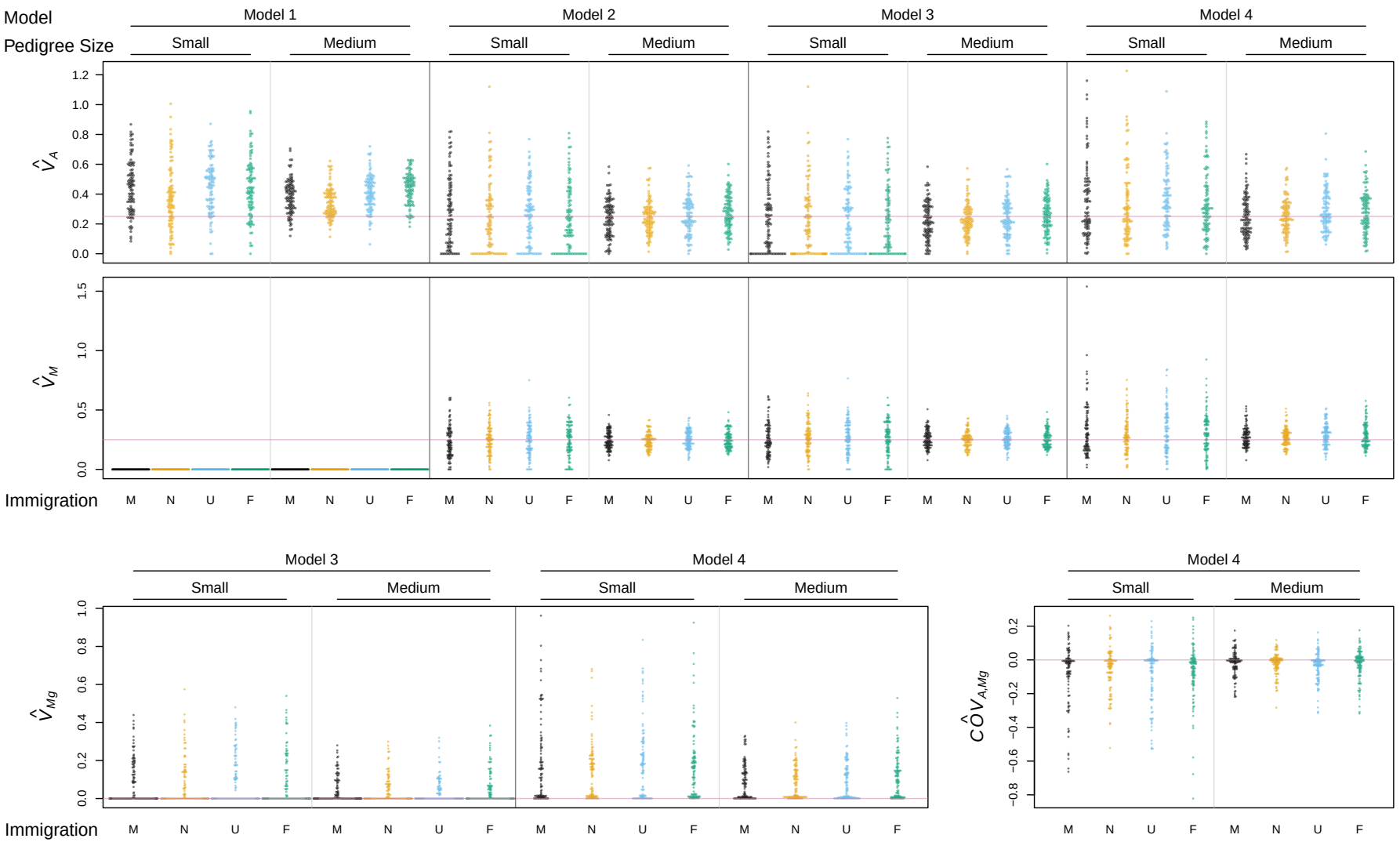


Figure S41: $\hat{V}_A$, $\hat{V}_M$, $\hat{V}_{Mg}$ and $C\hat{O}V_{A,Mg}$ from small and medium sized pedigree simulations across all simulated pedigrees from Scenario L (V_A =0.25, V_{Mg} =0, V_{Me} =0.25,$COV_{A,Mg}$ =0). M, N, U and F, stand for male-biased, no, unbiased and female-biased immigration respectively.