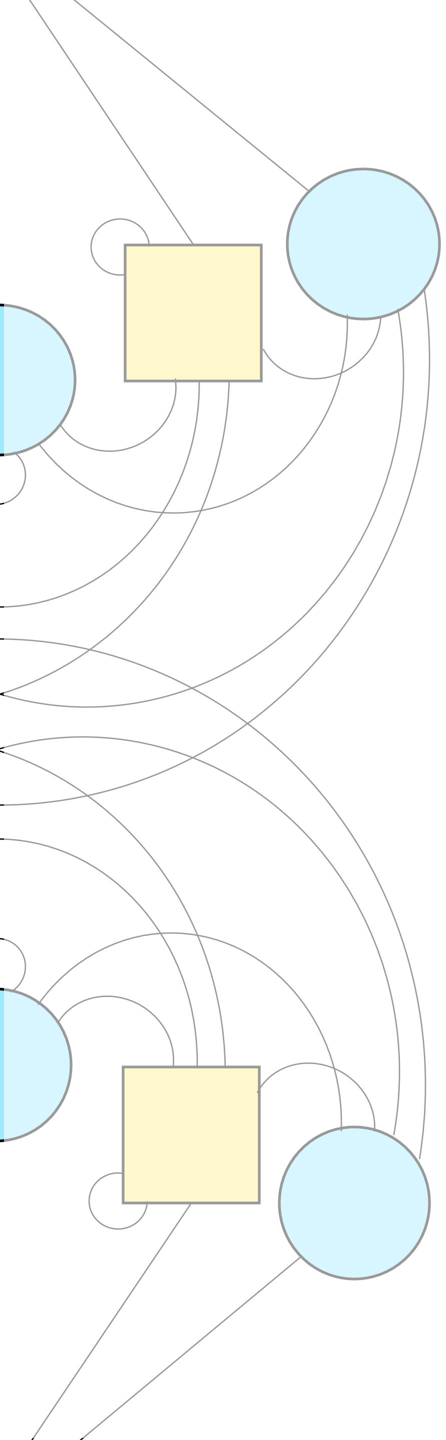
Estimating Parental Effects using Polygenic Scores Part II: Model Extensions

Jared V. Balbona | Yongkang Kim | Matthew C. Keller

Outline of this video:

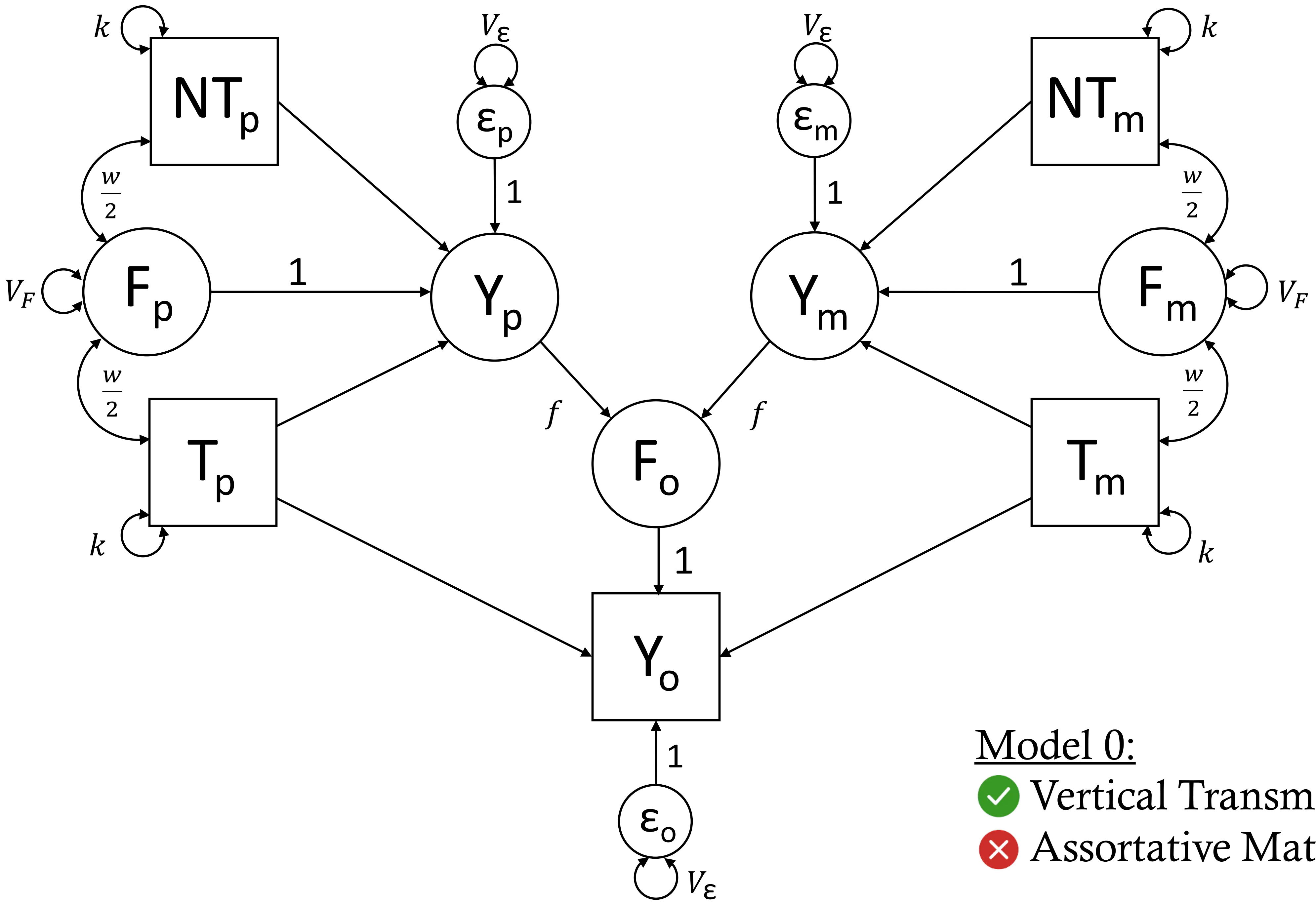
- How SEM-PGS models can detect and estimate different types of assortative mating
    - Disequilibrium vs. Equilibrium
    - Phenotypic, Social, and Genetic Homogamy
- Using latent genetic factors to capture full additive genetic effects
- Other model extensions
    - Different data types that can be leveraged
    - Different questions that can be addressed



Accounting for Assortative Mating

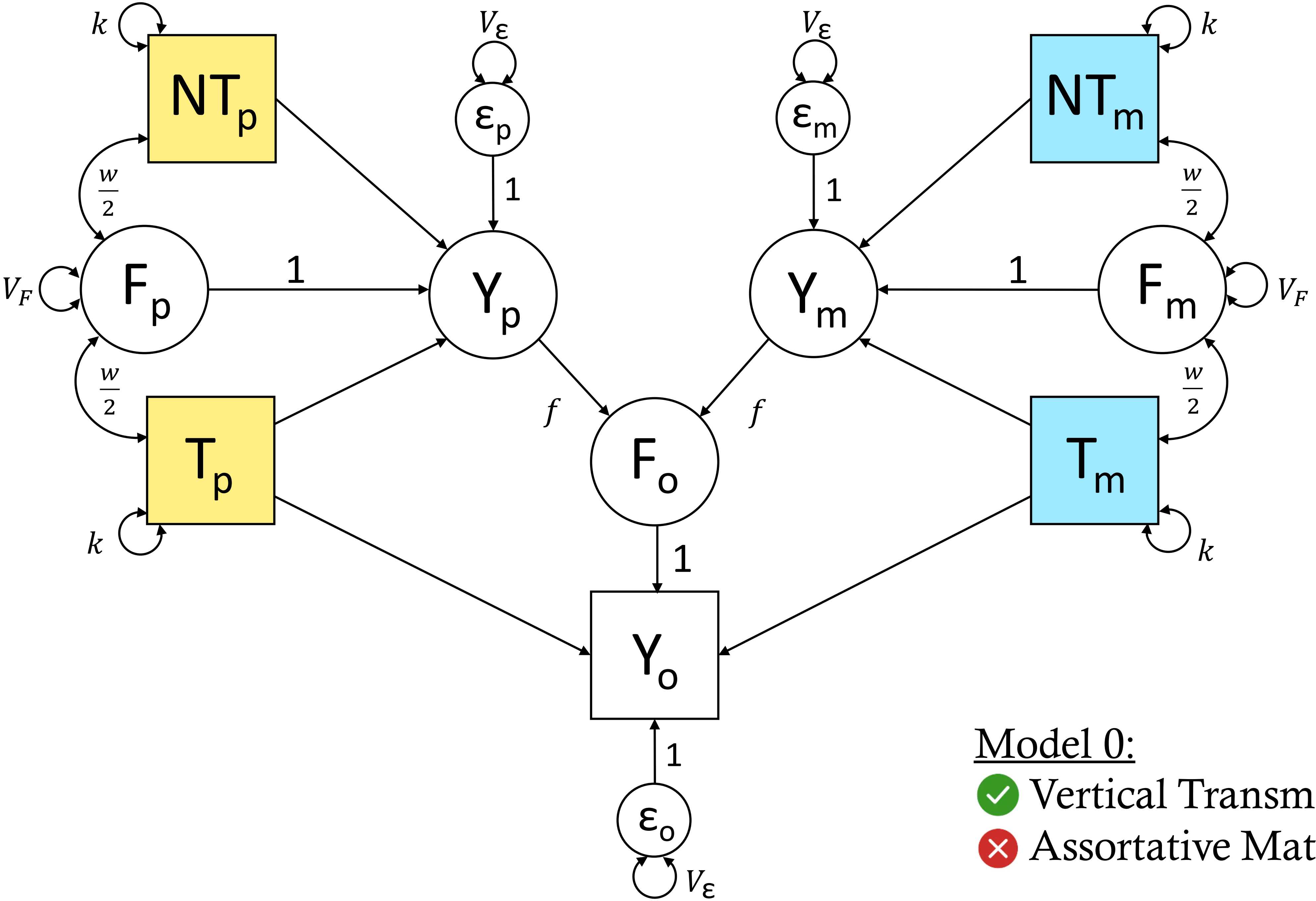
Importance of Assortative Mating

- Assortment is interesting in its own right!
  - Informs about a given population's culture, priorities, and preferences
- Assortment can impact genetic architecture
  - Assortative mating can lead to increases in:
    1. Correlations between trait-increasing alleles
       - *AKA "Gametic Phase Disequilibrium"*
    2. Genetic correlations between traits
    3. Homozygosity
    4. Genetic correlations between relatives
    5. Population phenotypic and genotypic variance



Model 0:

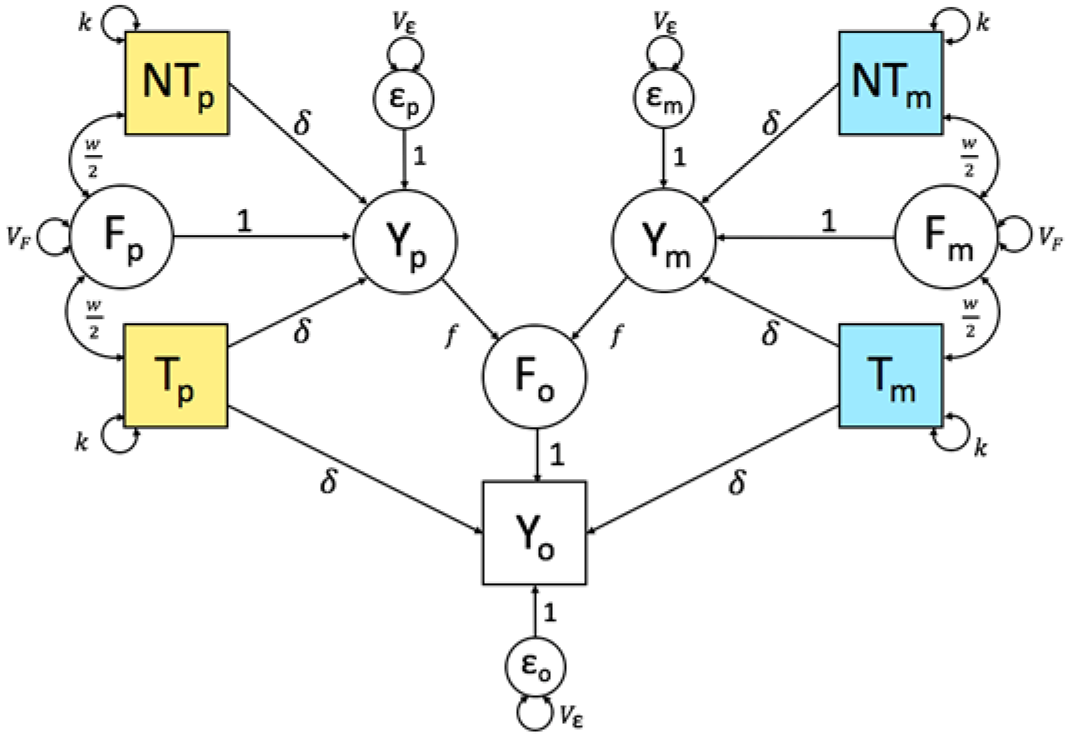
- ✓ Vertical Transmission
- ✗ Assortative Mating



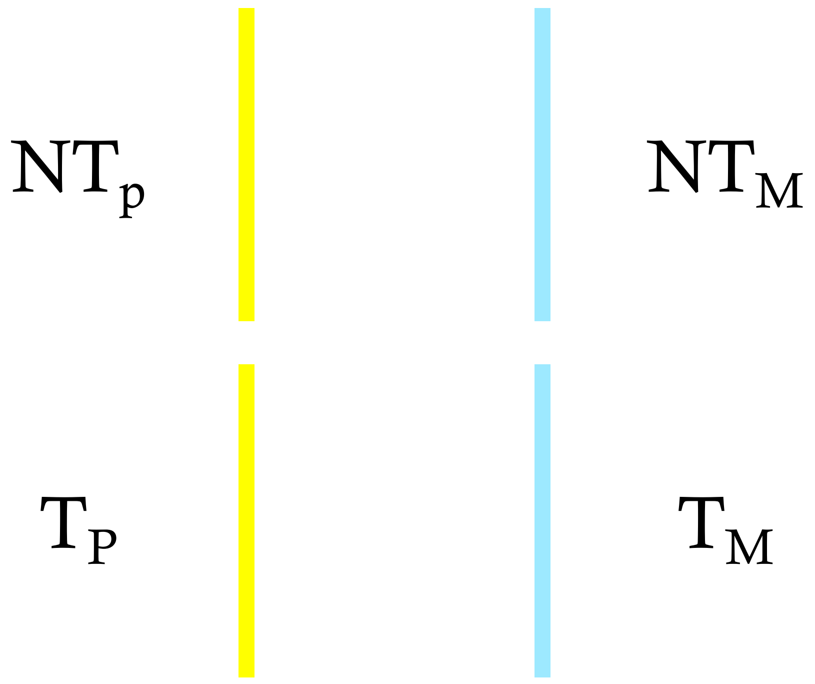
Model 0:

✓ Vertical Transmission

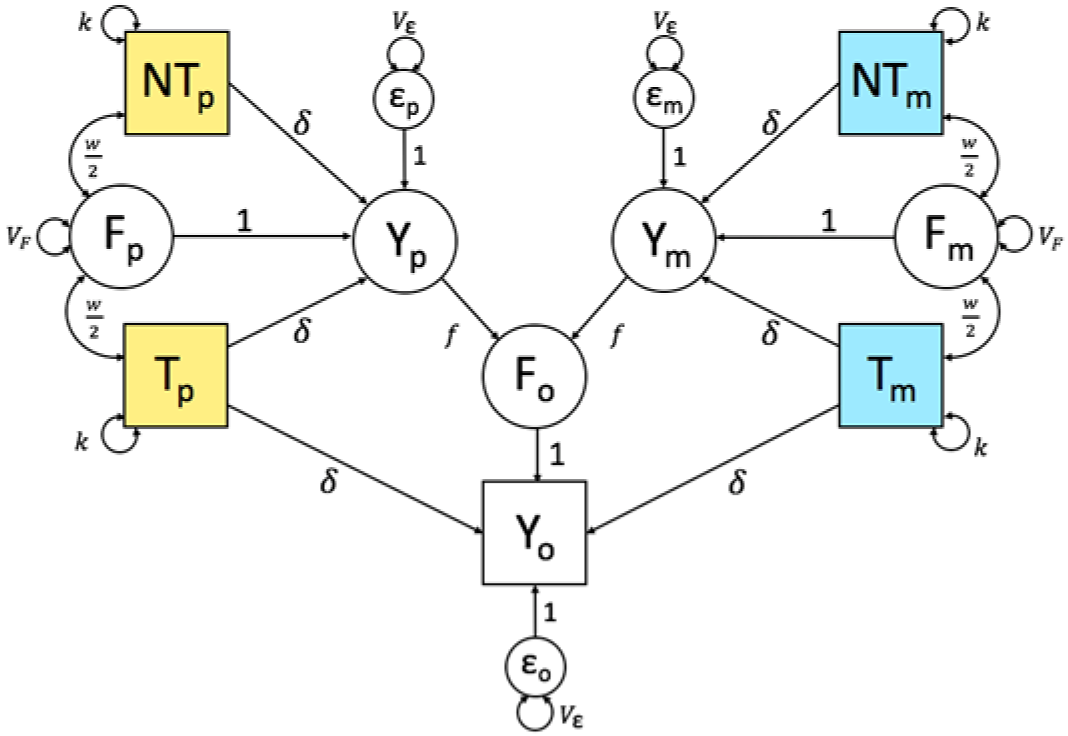
✗ Assortative Mating

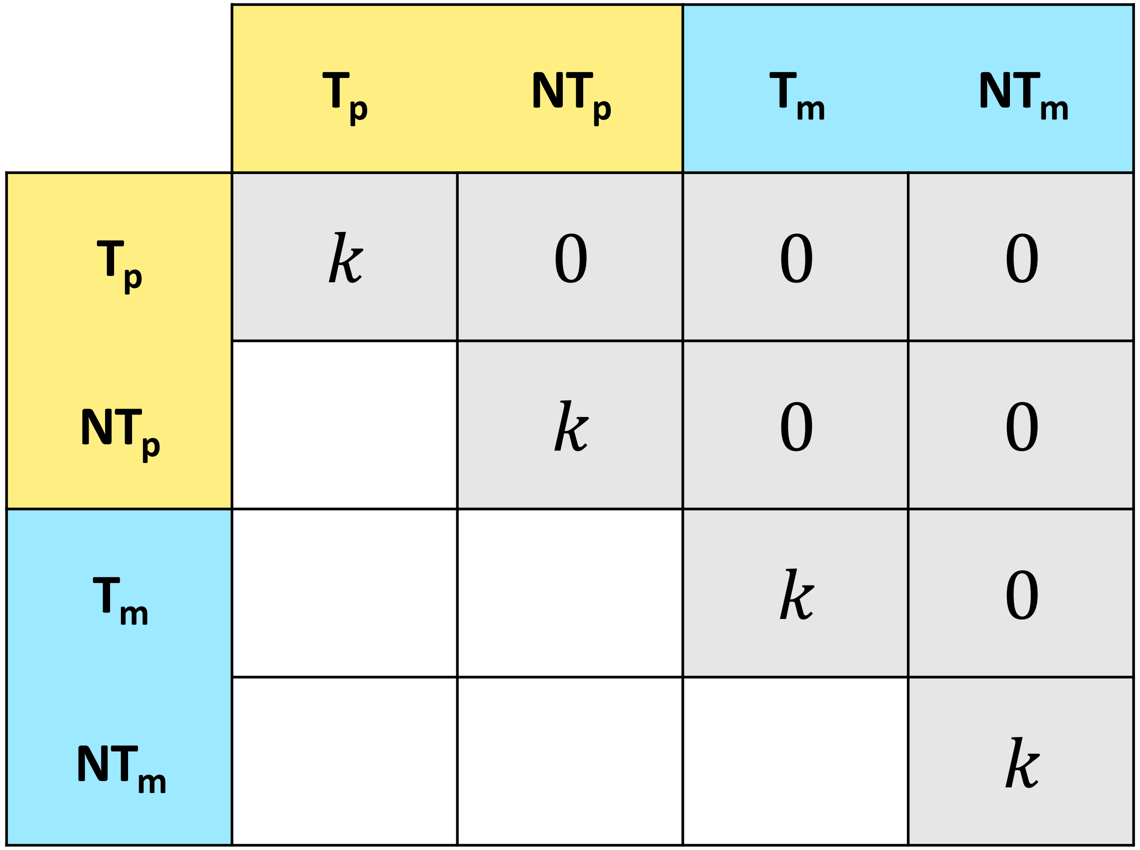


Offspring's Genome



No Assortment

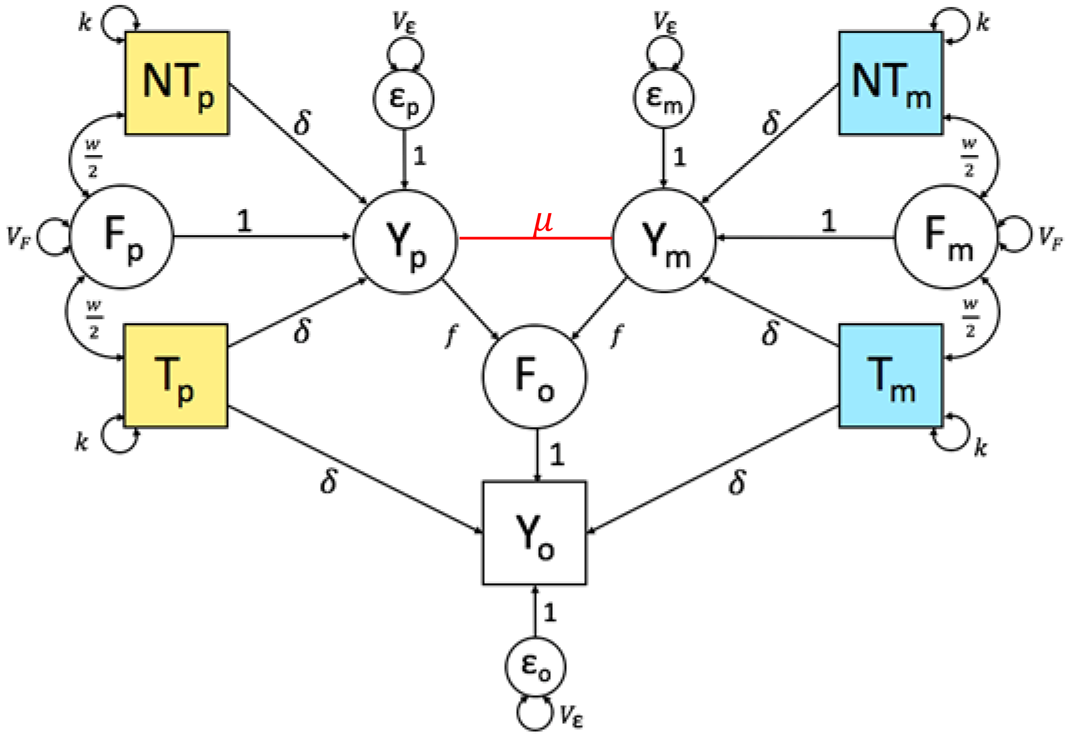




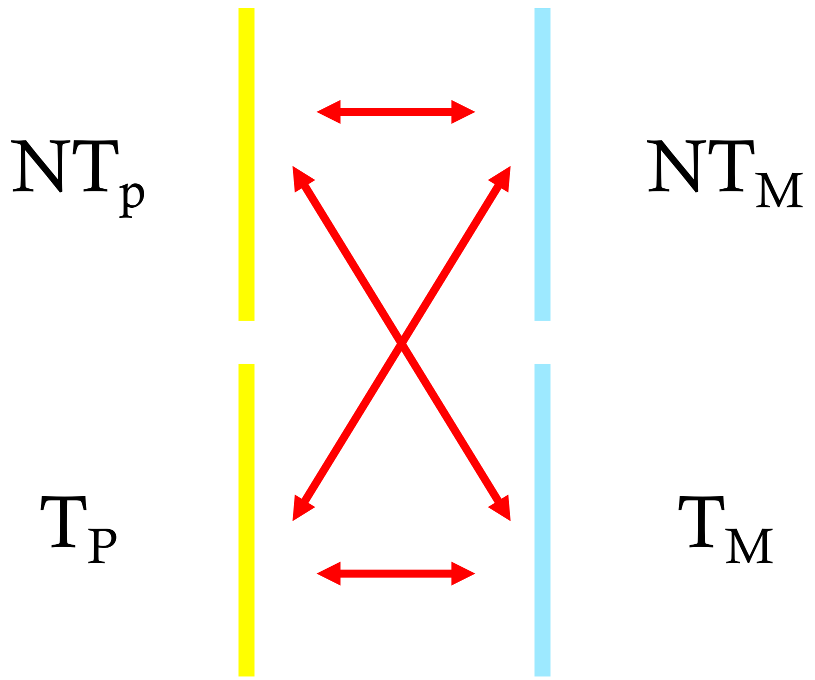
	T_p	NT_p	T_m	NT_m
T_p	k	0	0	0
NT_p		k	0	0
T_m			k	0
NT_m				k

No Assortment

$$\mu = \frac{\text{Spousal Phenotypic Covariance}}{(\text{Spouse Phenotypic Variance})^2}$$

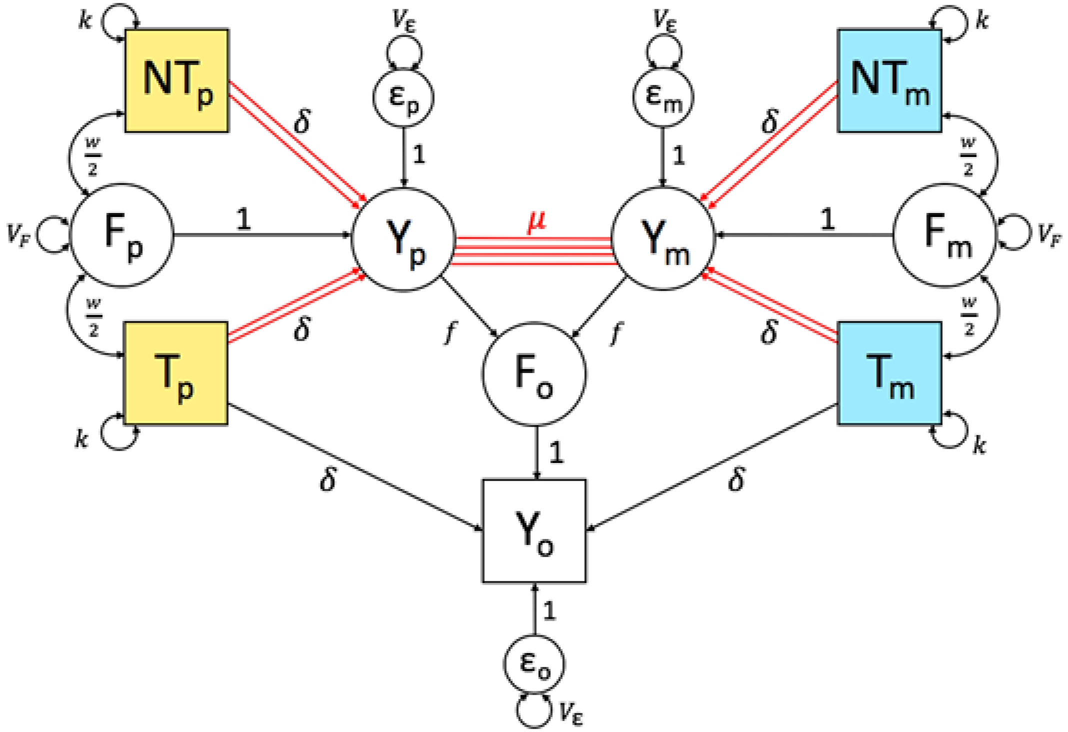


Offspring's Genome

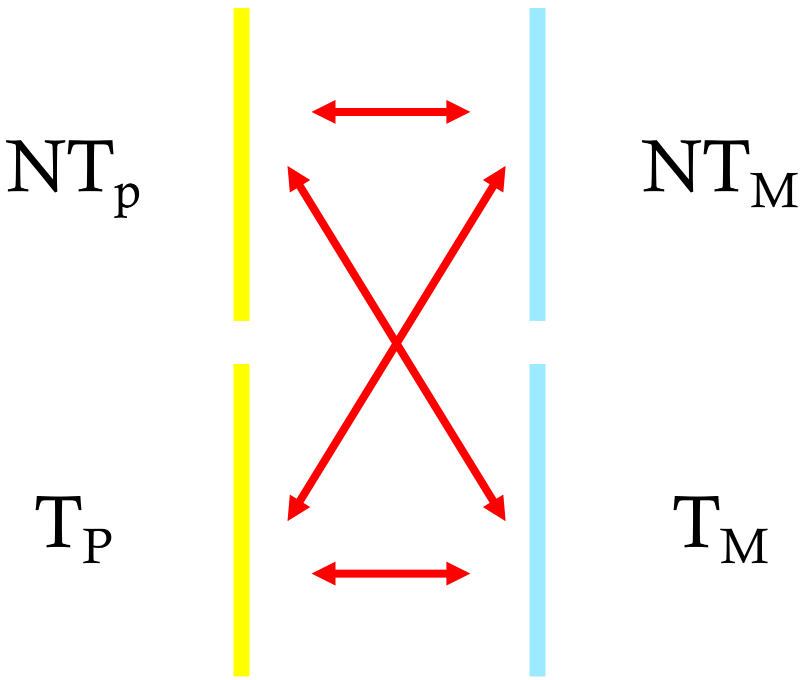


1 Generation of Assortment

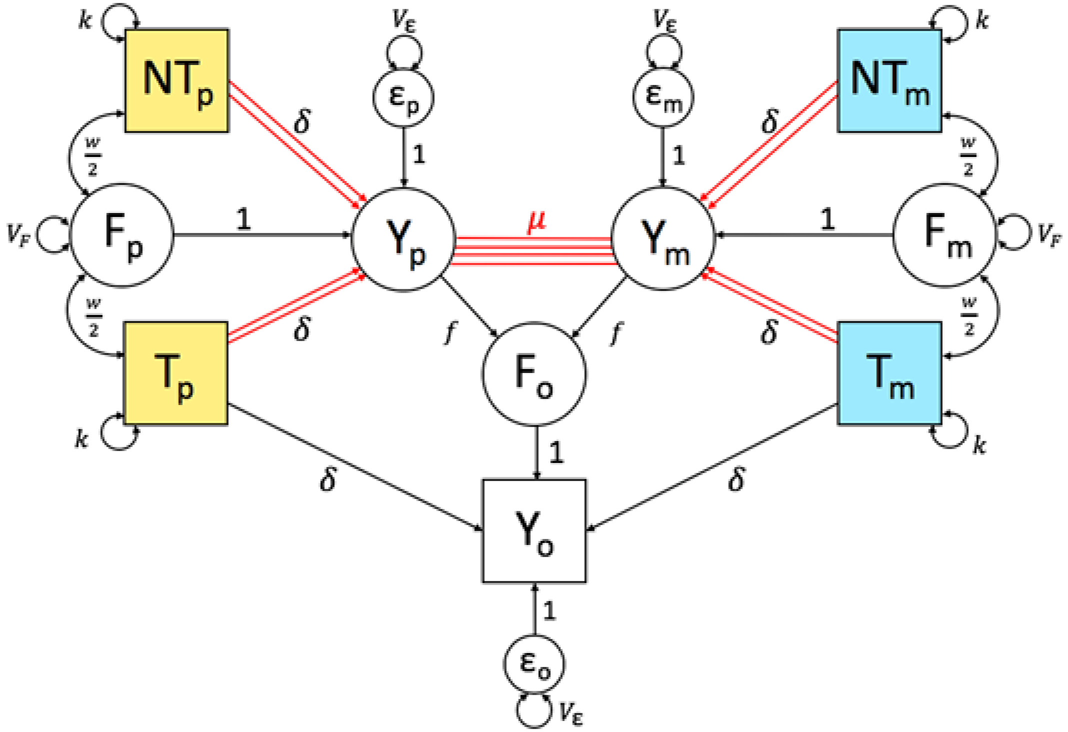
$$\mu = \frac{\text{Spousal Phenotypic Covariance}}{(\text{Spouse Phenotypic Variance})^2}$$



Offspring's Genome

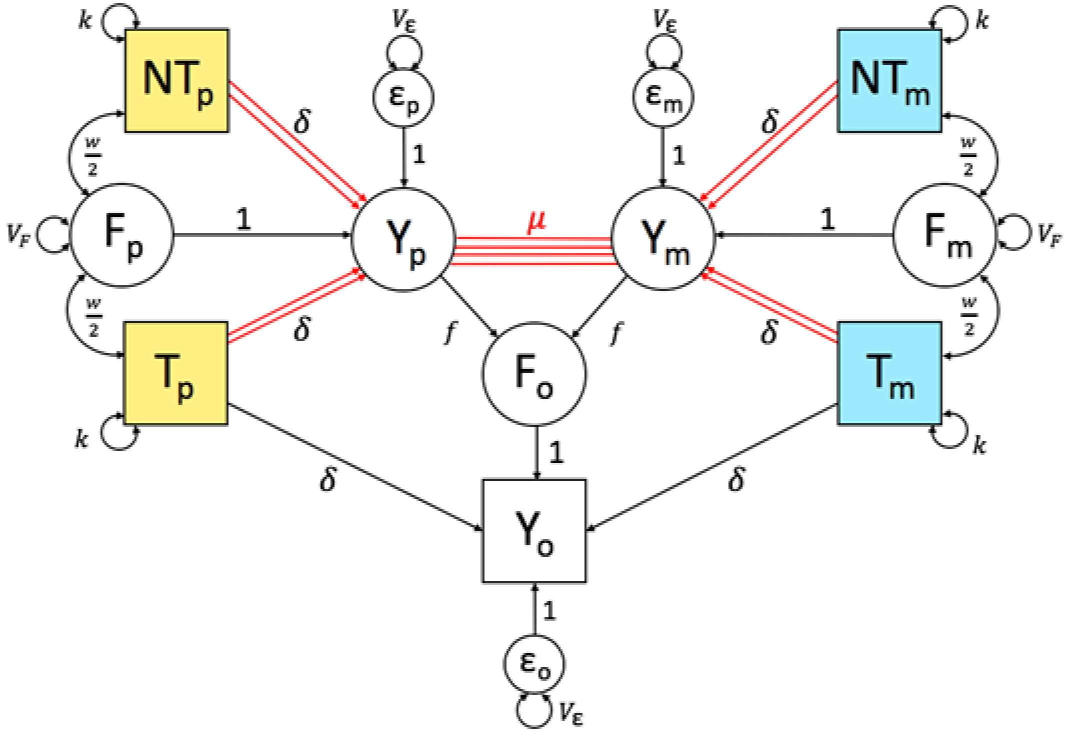


1 Generation of Assortment



	T_p	NT_p	T_m	NT_m
T_p	k	0	0	0
NT_p		k	0	0
T_m			k	0
NT_m				k

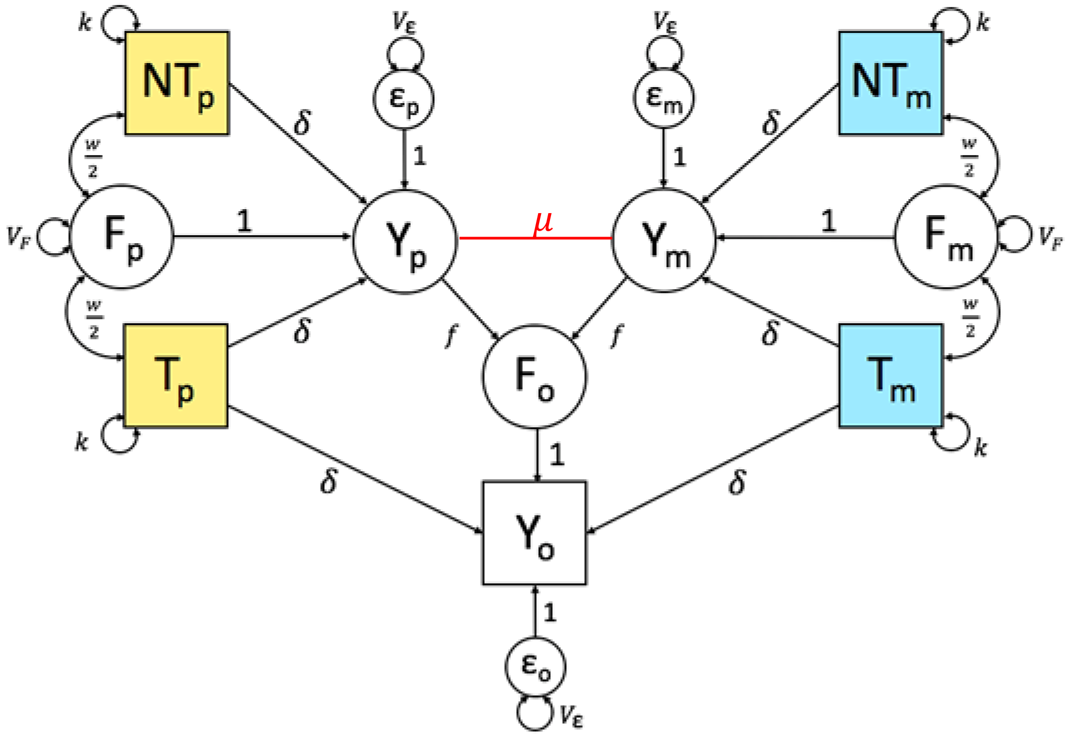
1 Generation of Assortment



g : Increase in PGS covariances due to AM

	T_p	NT_p	T_m	NT_m
T_p	k	0	g	g
NT_p		k	g	g
T_m			k	0
NT_m				k

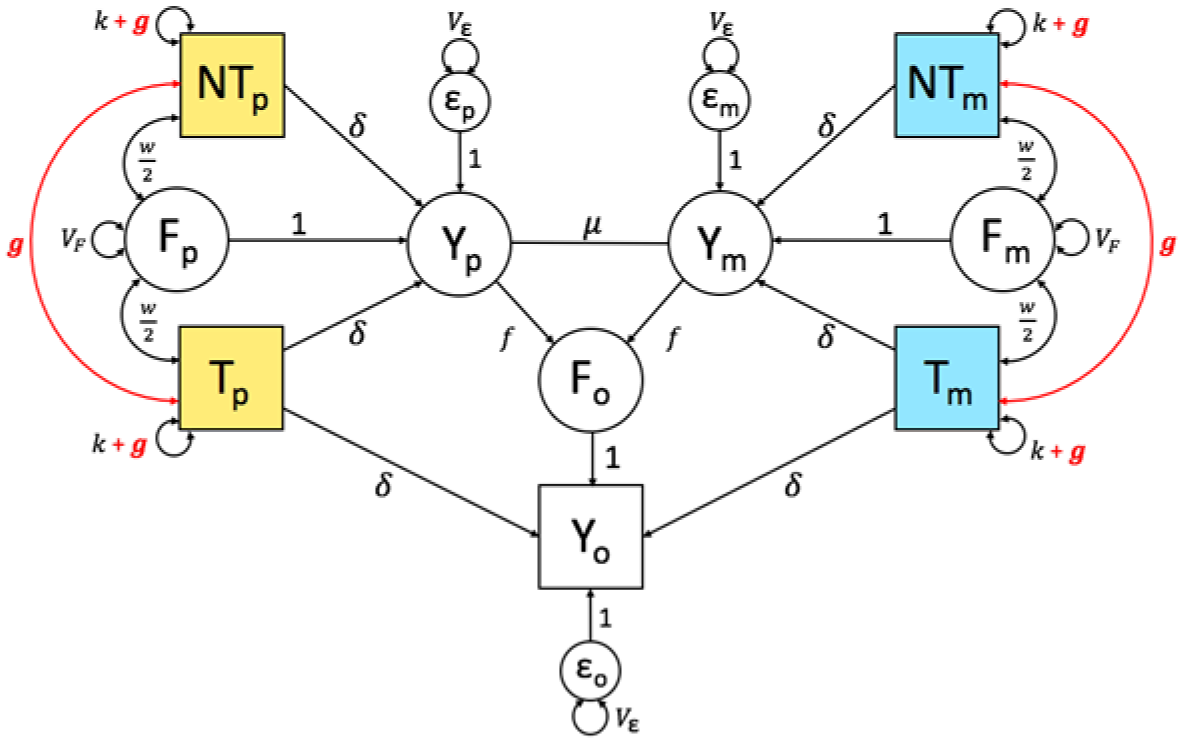
1 Generation of Assortment



g : Increase in PGS covariances due to AM

	T_p	NT_p	T_m	NT_m
T_p	k	0	g	g
NT_p		k	g	g
T_m			k	0
NT_m				k

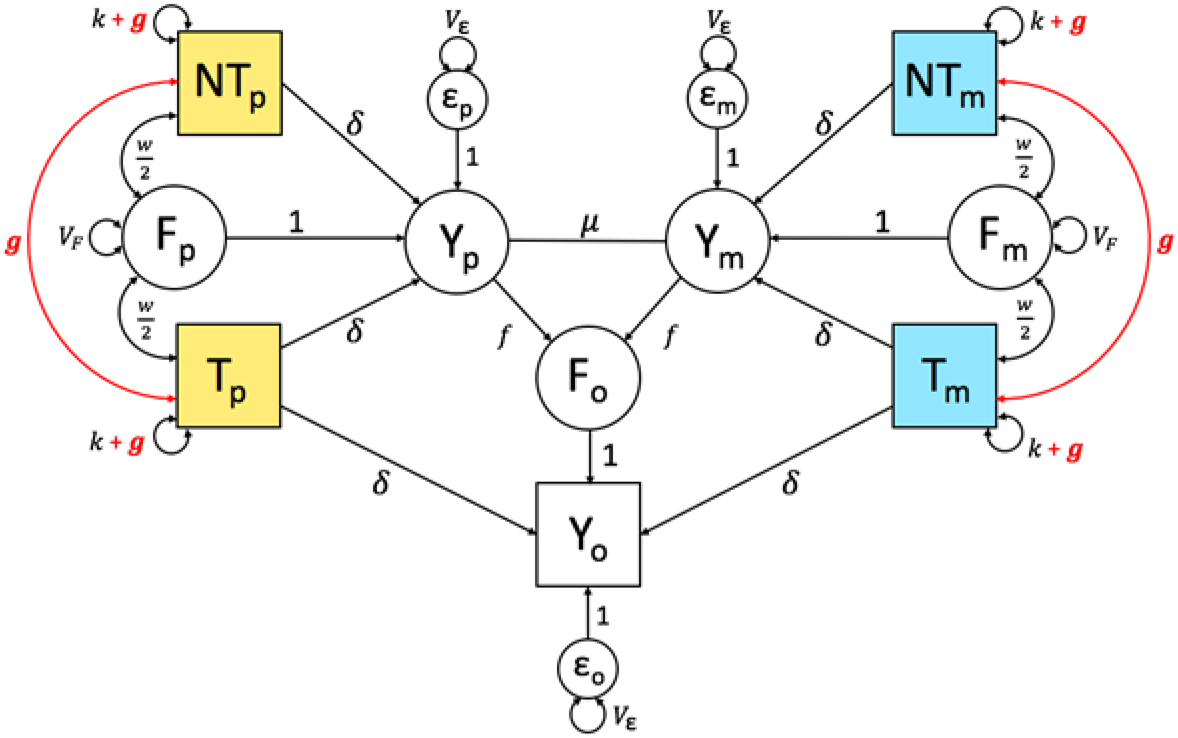
>1 Generation of Assortment



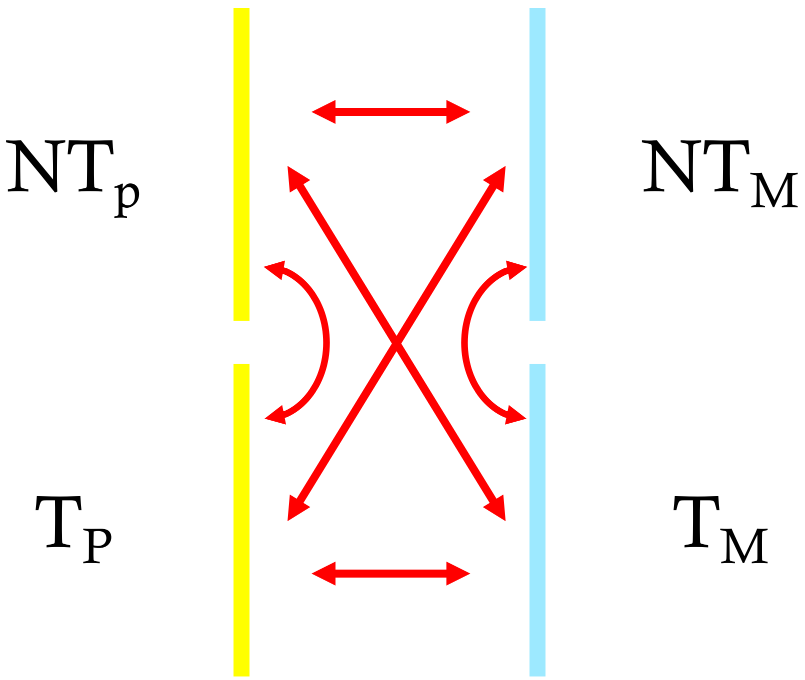
g : Increase in PGS (co)variances due to AM

	T_p	NT_p	T_m	NT_m
T_p	$k + g$	g	g	g
NT_p		$k + g$	g	g
T_m			$k + g$	g
NT_m				$k + g$

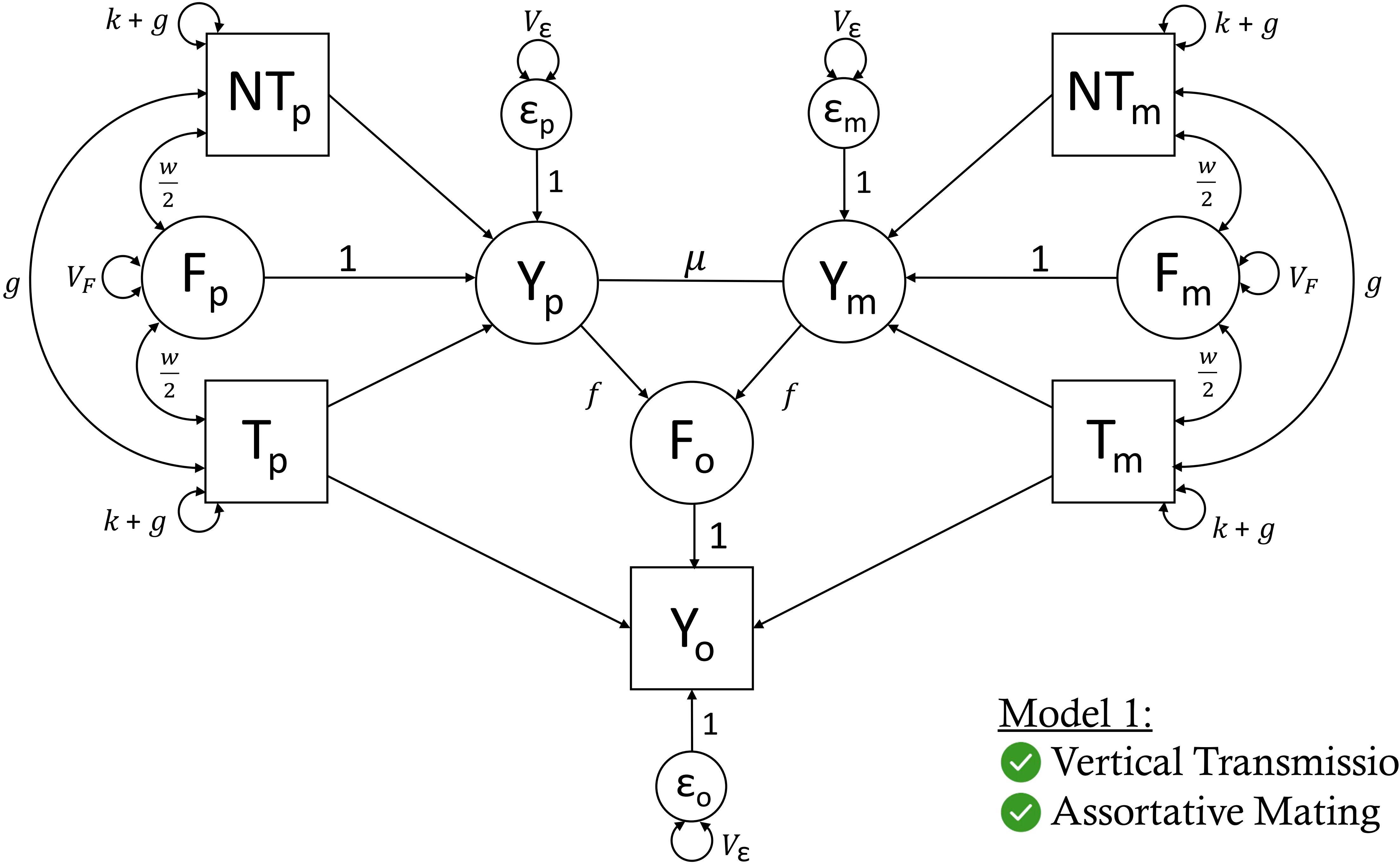
>1 Generation of Assortment



Offspring's Genome



>1 Generation of Assortment



Model 1:

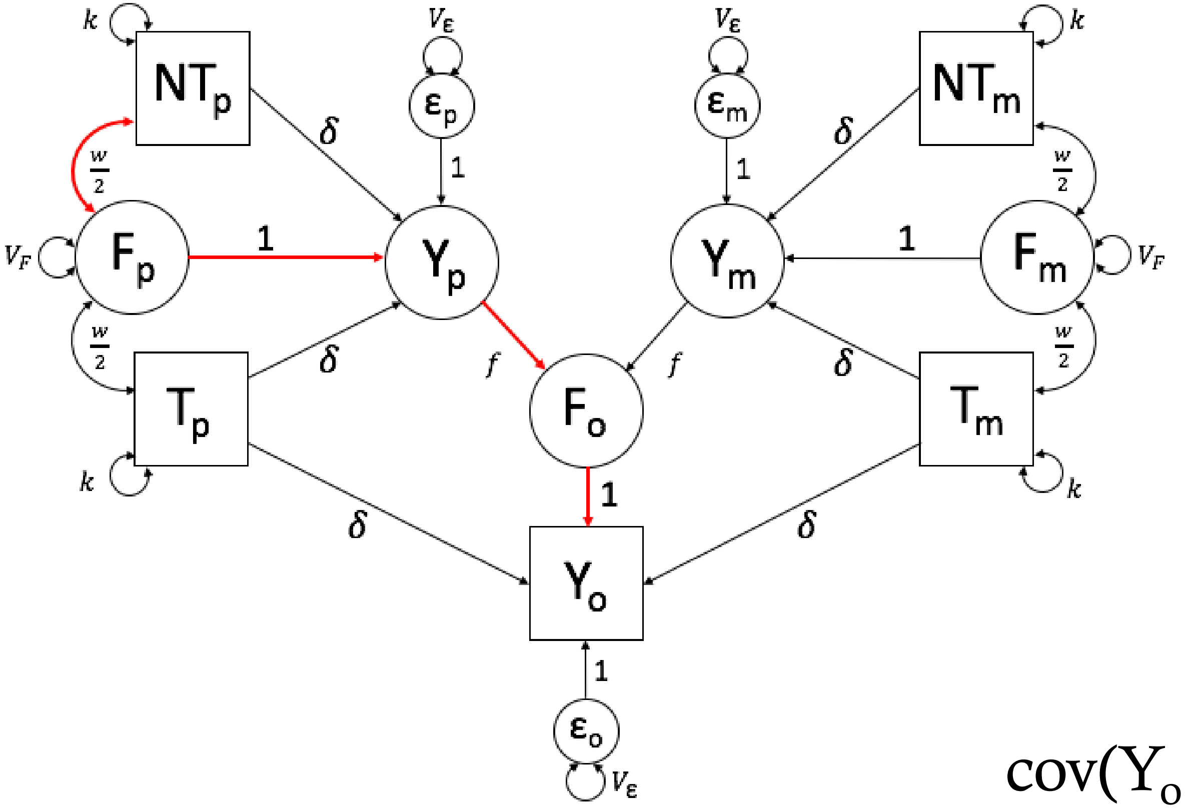
✓ Vertical Transmission

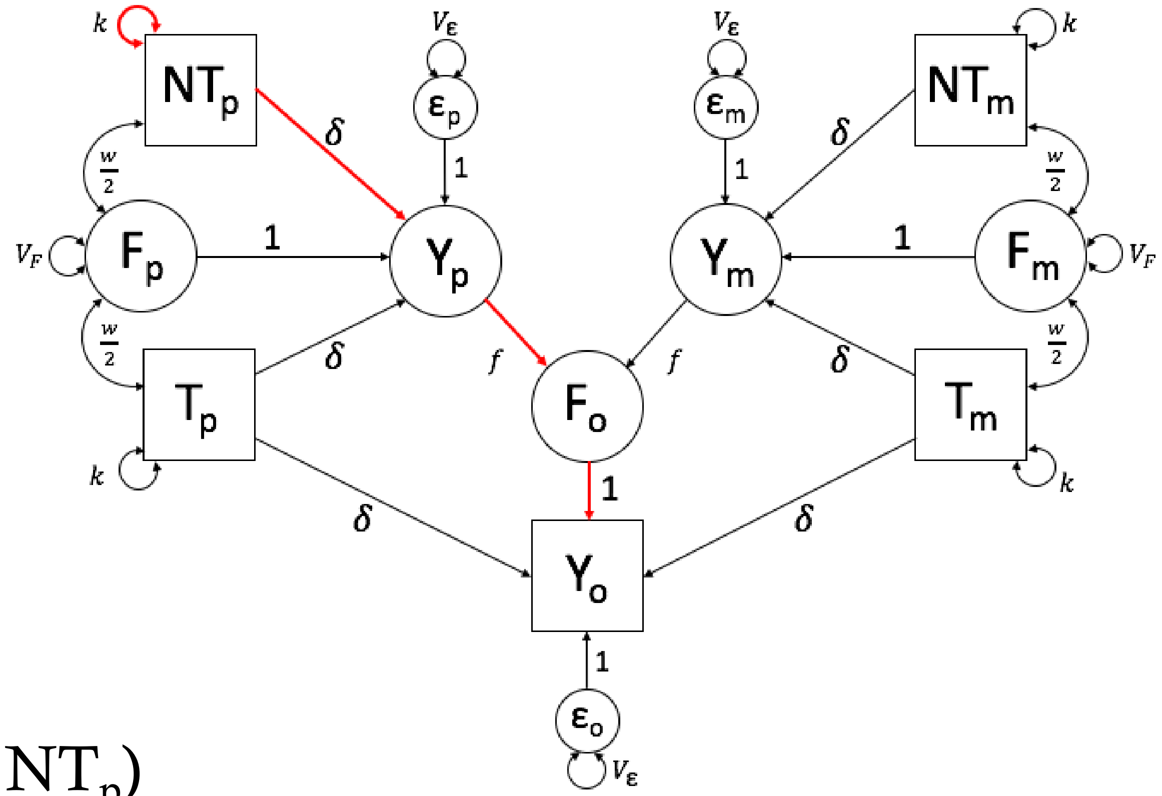
✓ Assortative Mating

Model 0:

✓ Vertical Transmission,

✗ Assortative Mating





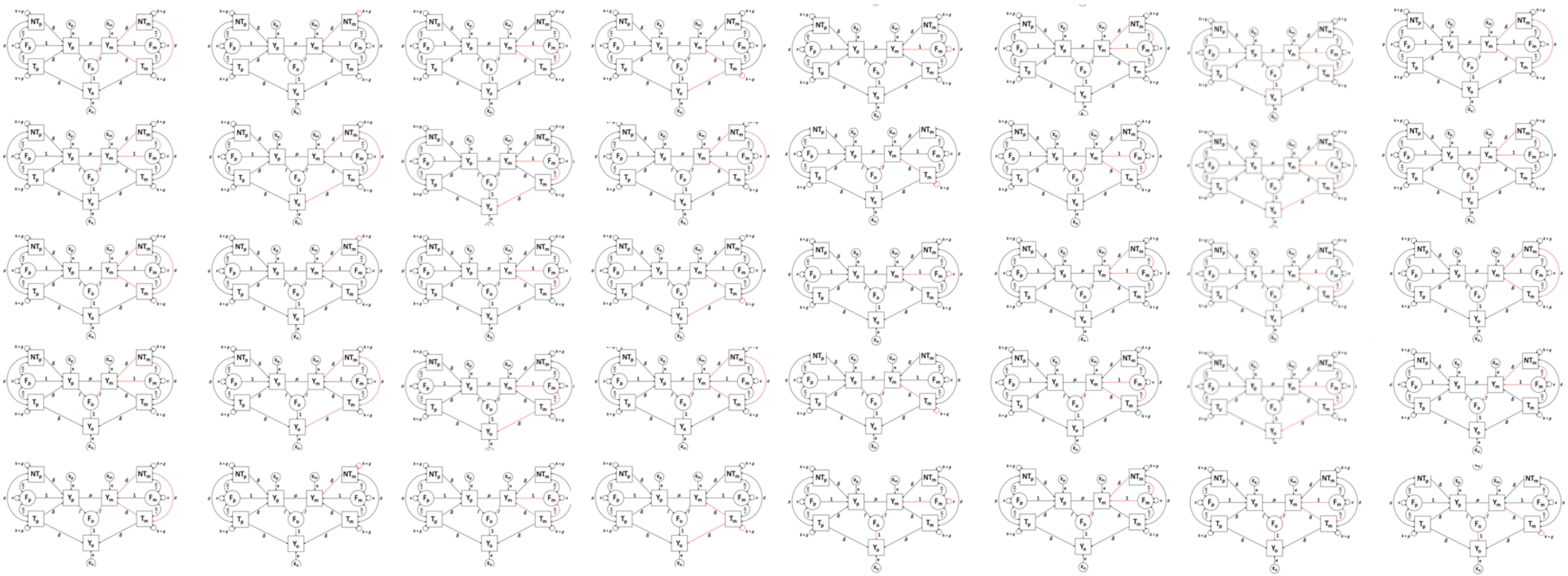
$$\mathrm{cov}(\mathrm{Y_o}\ , \mathrm{NT_p})$$

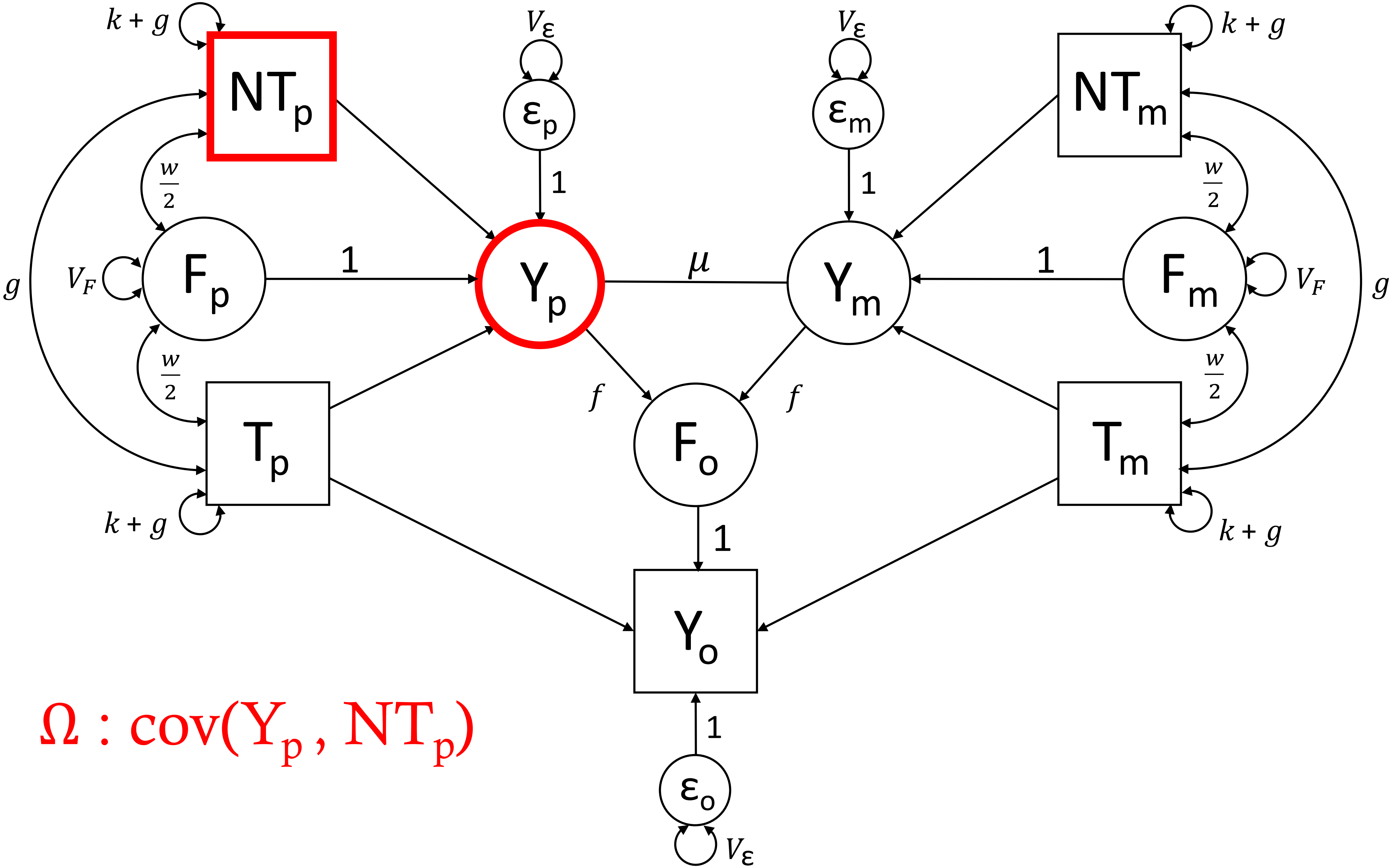
$$= \left(1 * f * 1 * \tfrac{w}{2}\right) + \left(1 * f * \delta * k\right)$$

$$= \tfrac{1}{2} f w + f \delta k$$

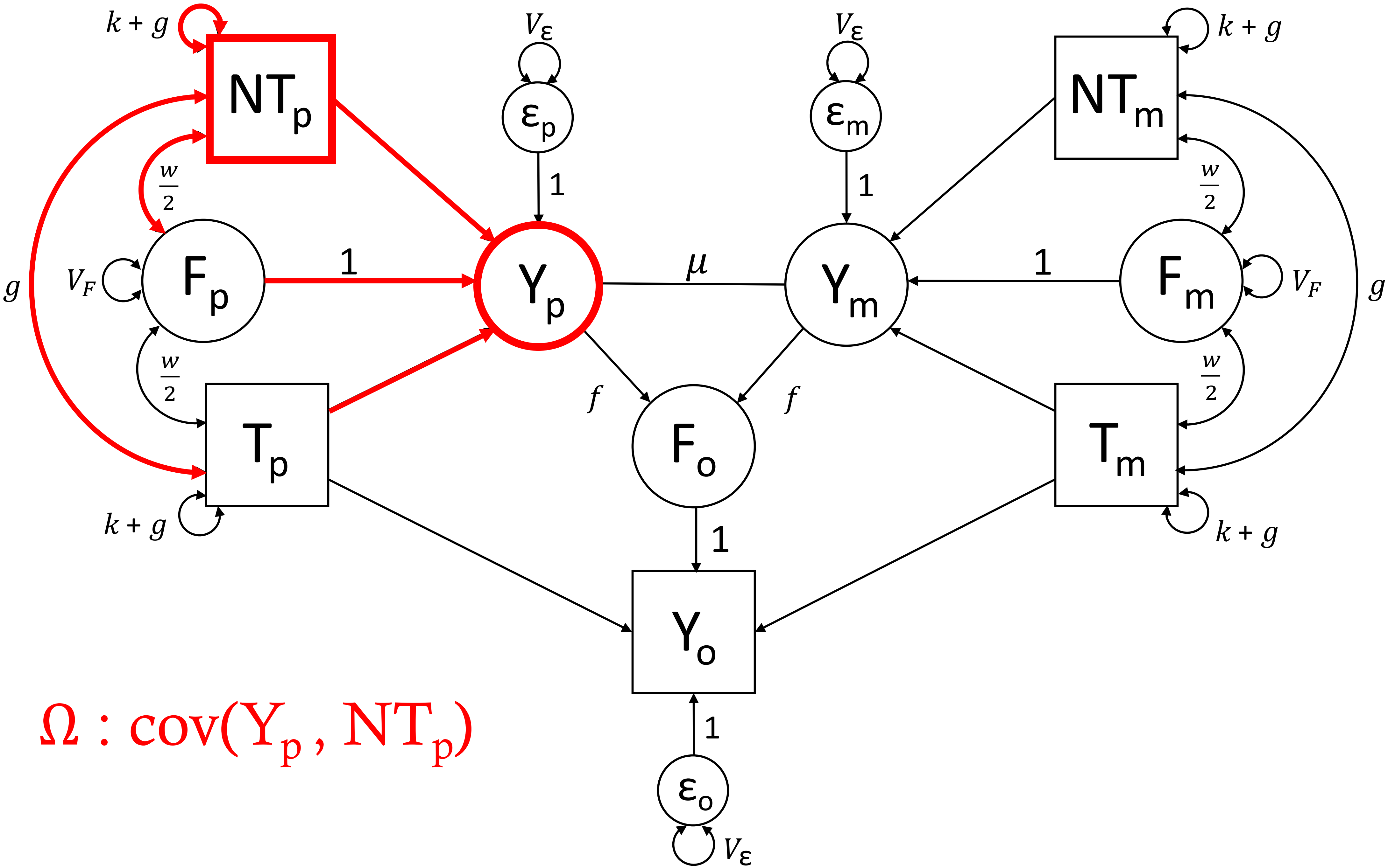
Model 1:

- ✓ Vertical Transmission,
- ✓ Assortative Mating





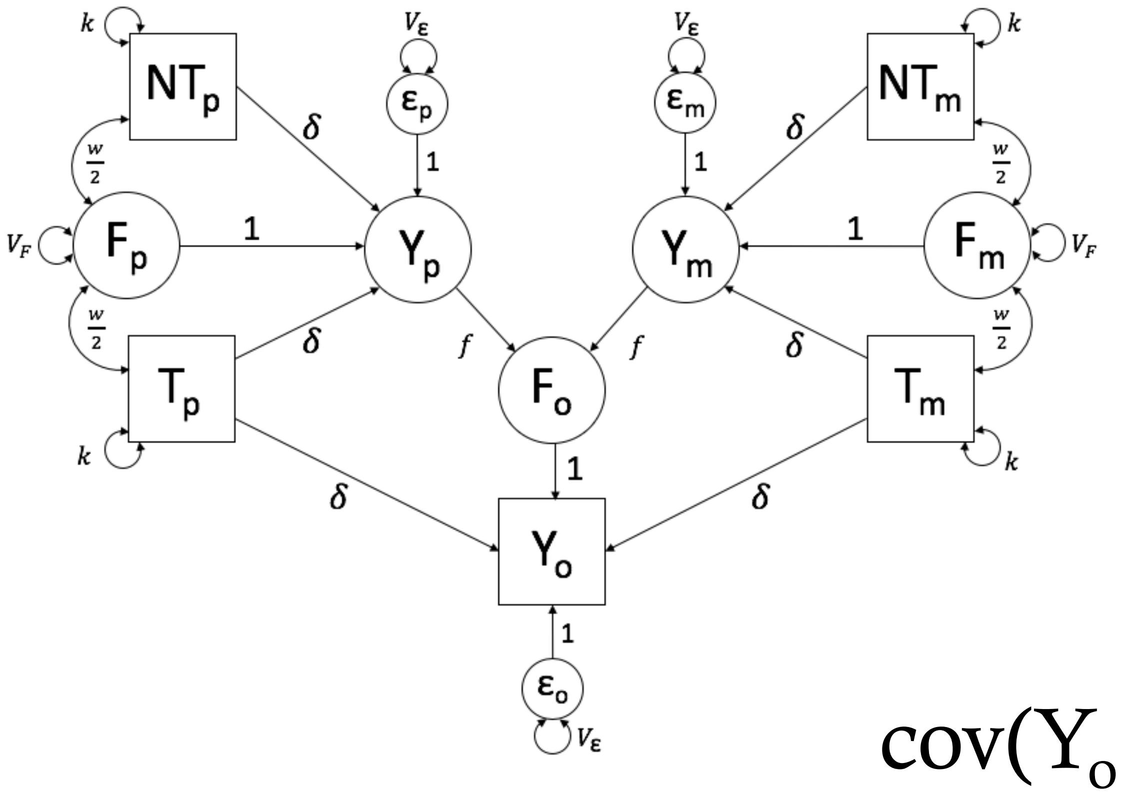
$\Omega : \mathrm{cov}(Y_p, NT_p)$



$\Omega : \mathrm{cov}(Y_p\,, NT_p)$

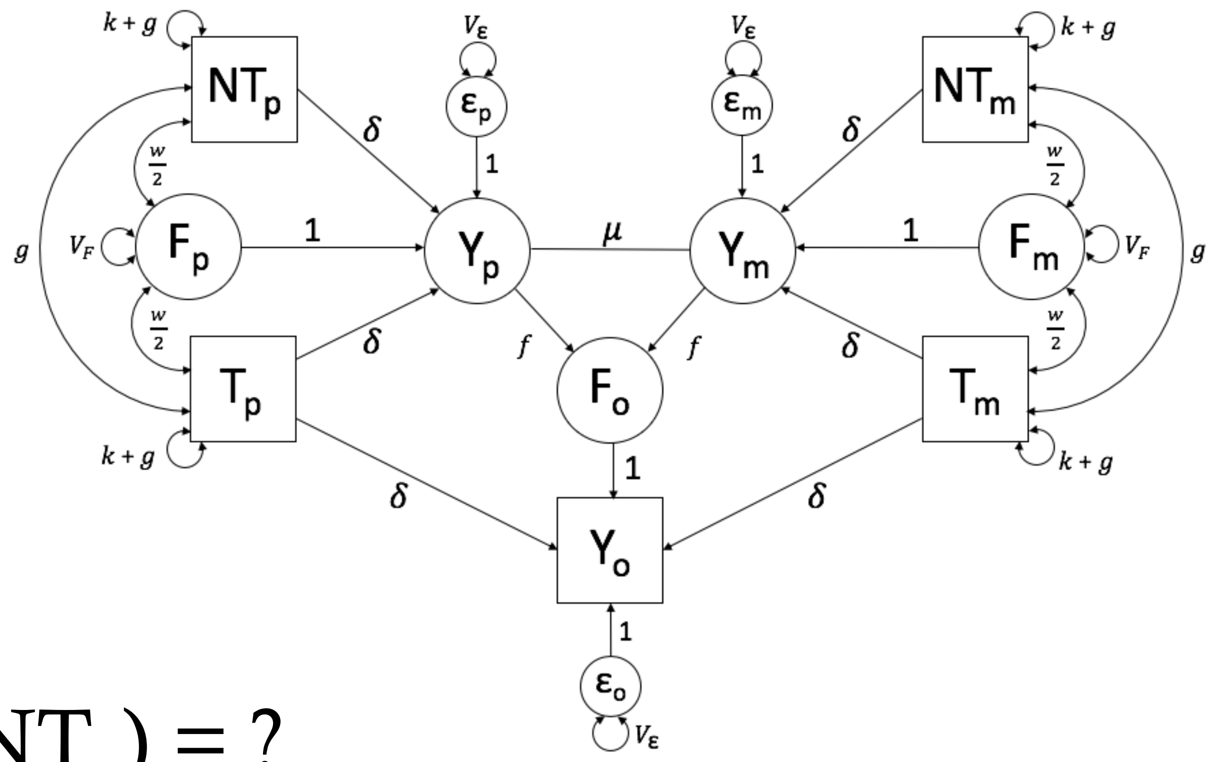
Model 0:

✓ Vertical Transmission,
✗ Assortative Mating



Model 1:

✓ Vertical Transmission,
✓ Assortative Mating



$$\mathrm{cov}(\mathrm{Y_o}\,,\,\mathrm{NT_p}) = ?$$

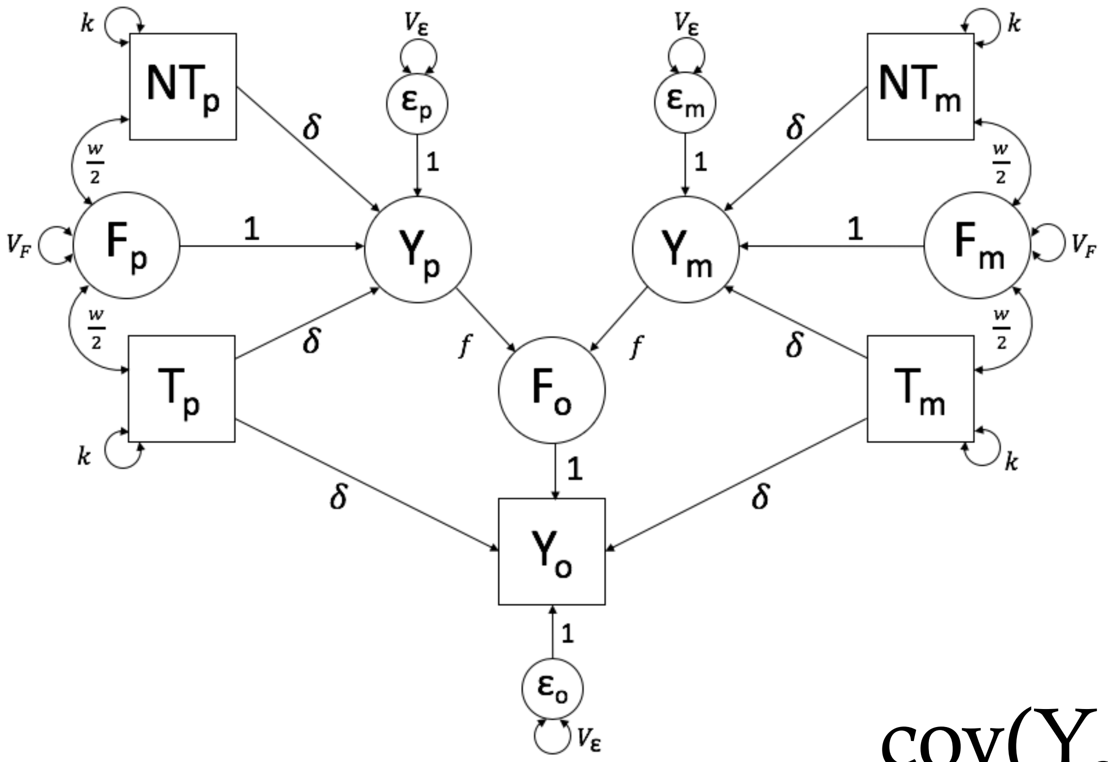
$$= f\left(\frac{w}{2} + \delta \mathrm{k}\right)$$

$$= f\Omega$$

$$= f\Omega\left(1 + V_y\mu\right) + 2\delta g$$

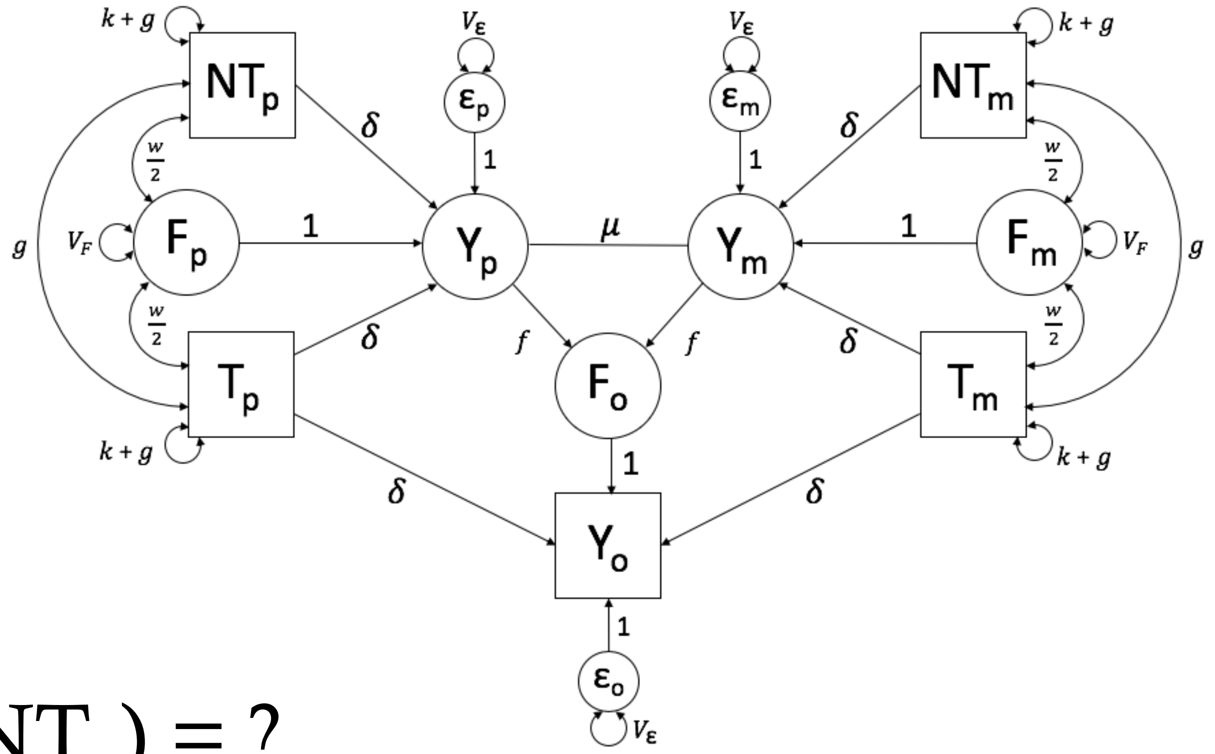
Model 0:

✓ Vertical Transmission,
✗ Assortative Mating



Model 1:

✓ Vertical Transmission,
✓ Assortative Mating



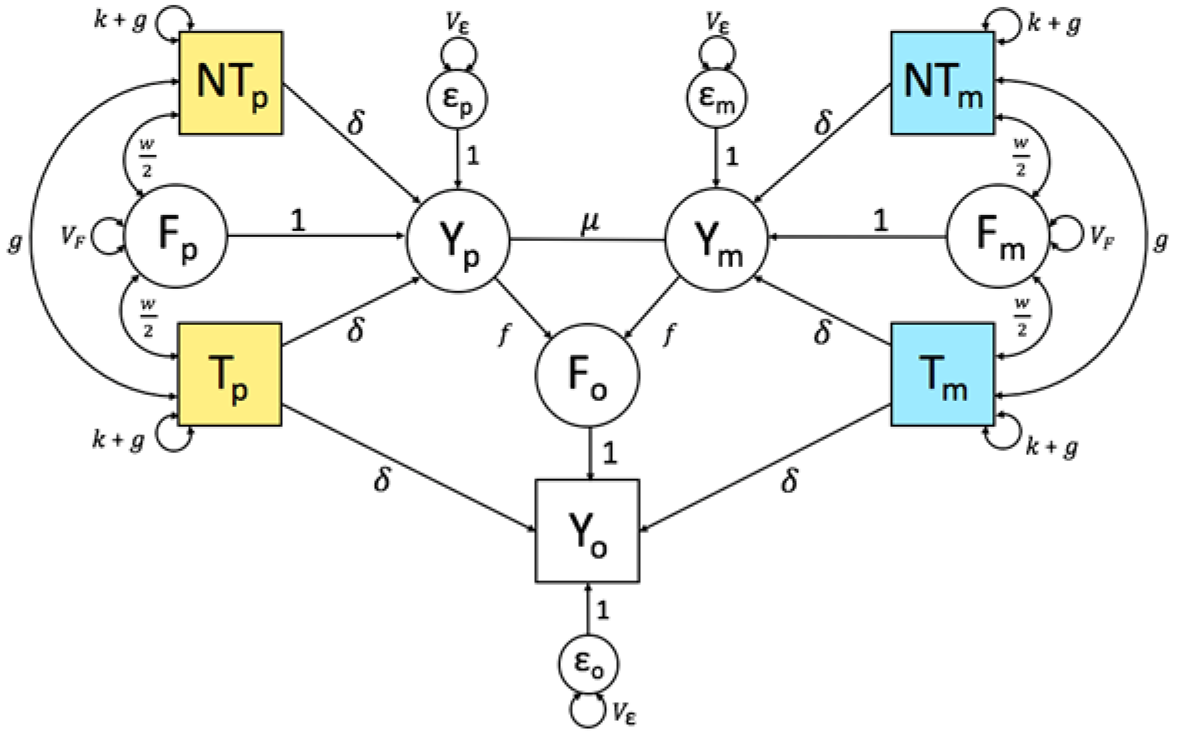
$$\mathrm{cov}(\mathrm{Y_o}\,,\,\mathrm{NT_p}) = \,?$$

$$= f\left(\frac{w}{2} + \delta \mathrm{k}\right)$$

$$= f\Omega$$

$$= f\Omega\big(1 + V_y\mu\big) + 2\delta g$$

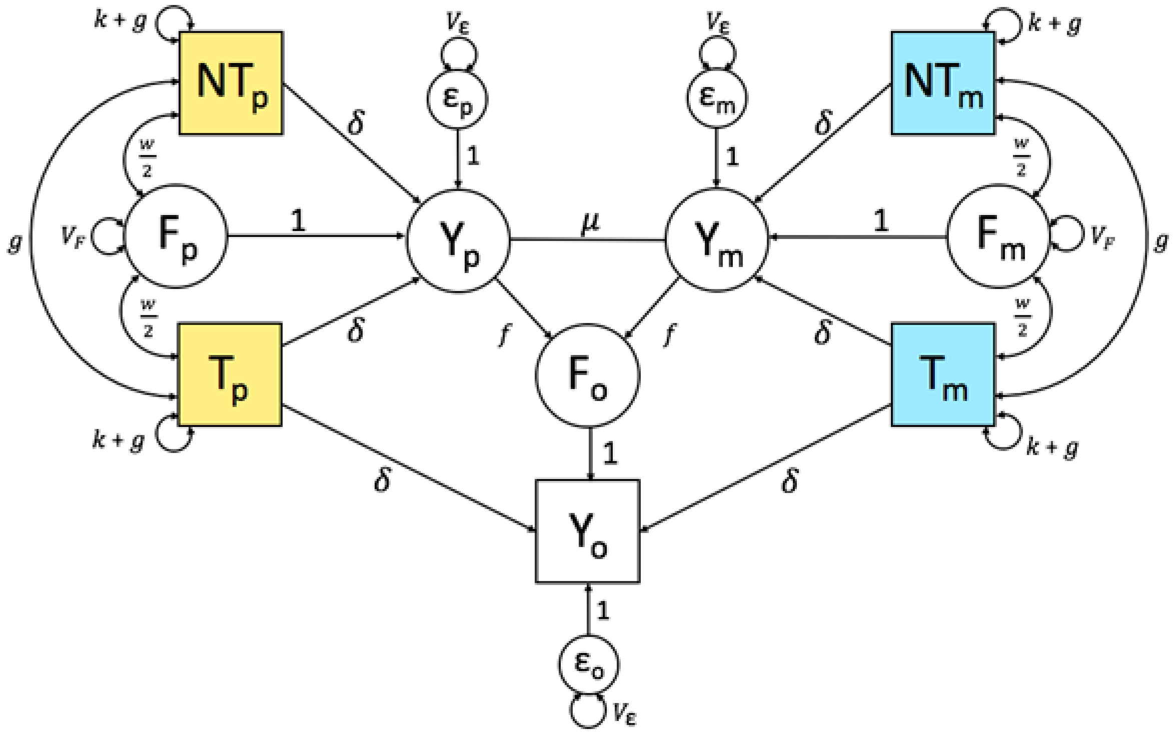
Determining if AM is changing over time

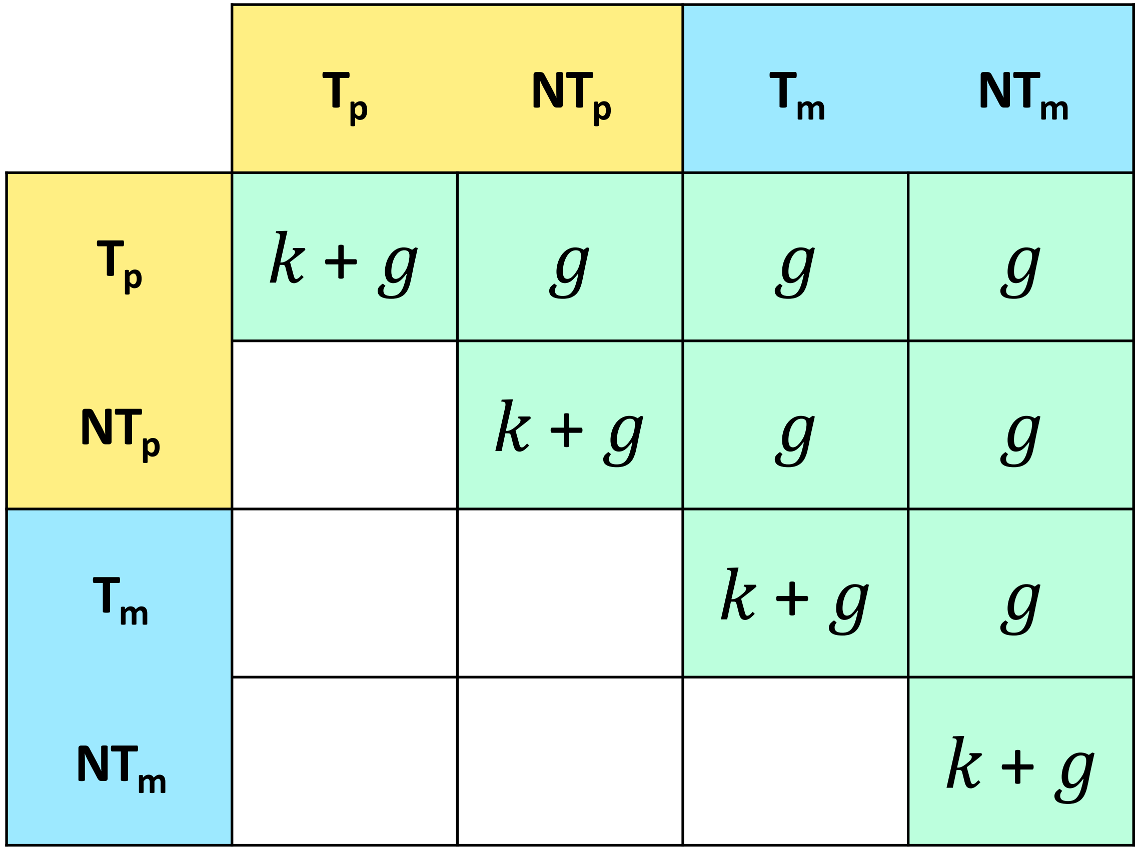


	$\mathrm{T_p}$	$\mathrm{NT_p}$	$\mathrm{T_m}$	$\mathrm{NT_m}$
$\mathrm{T_p}$	$k + g$	g	g	g
$\mathrm{NT_p}$		$k + g$	g	g
$\mathrm{T_m}$			$k + g$	g
$\mathrm{NT_m}$				$k + g$

Model 1:

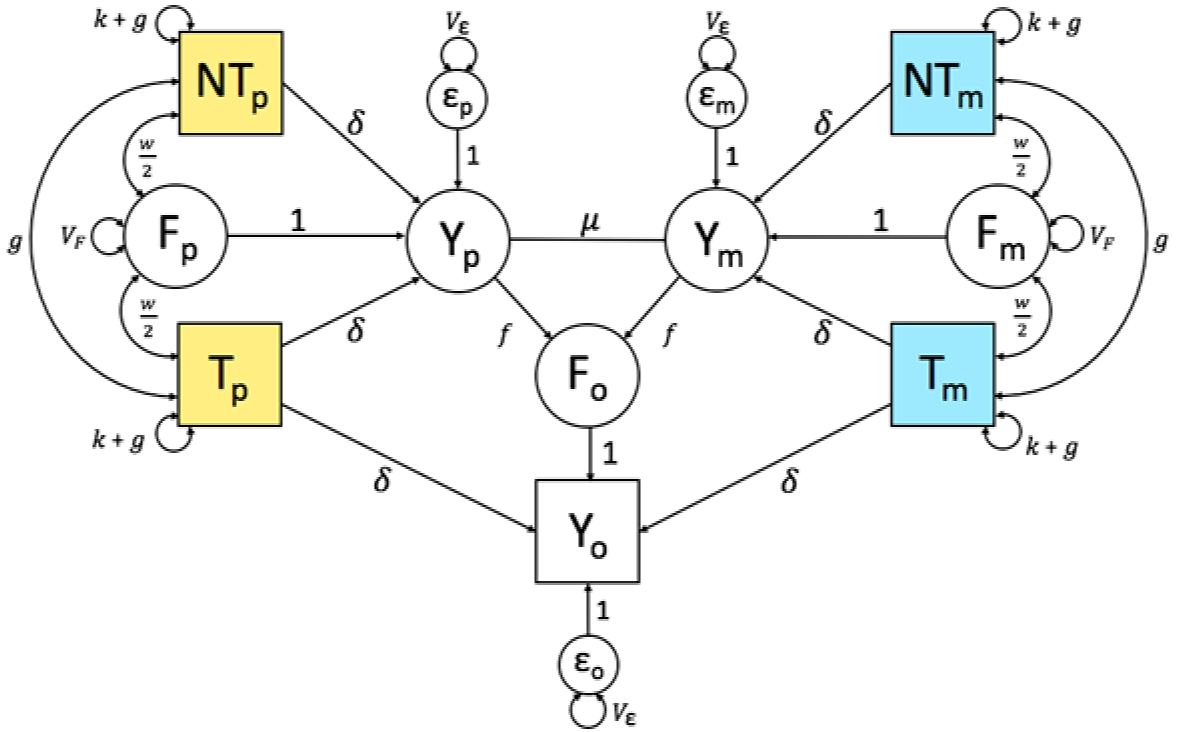
- ✓ Vertical Transmission
- ✓ Assortative Mating





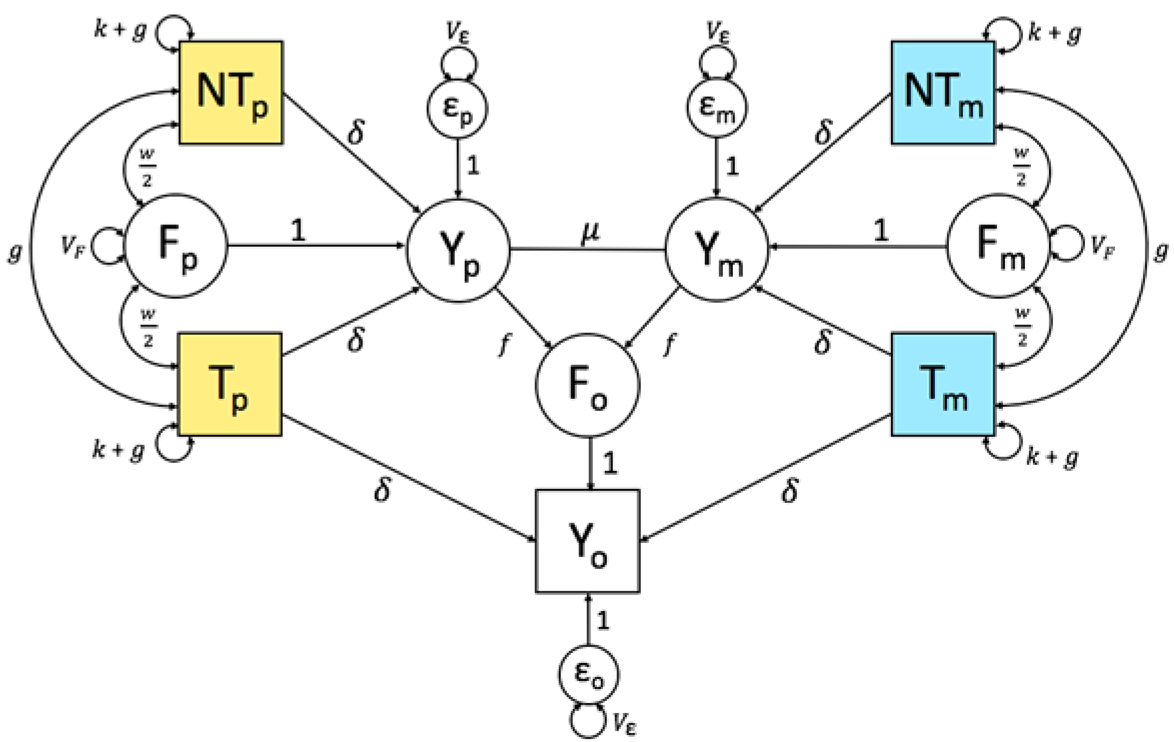
	T_p	NT_p	T_m	NT_m
T_p	$k+g$	g	g	g
NT_p		$k+g$	g	g
T_m			$k+g$	g
NT_m				$k+g$

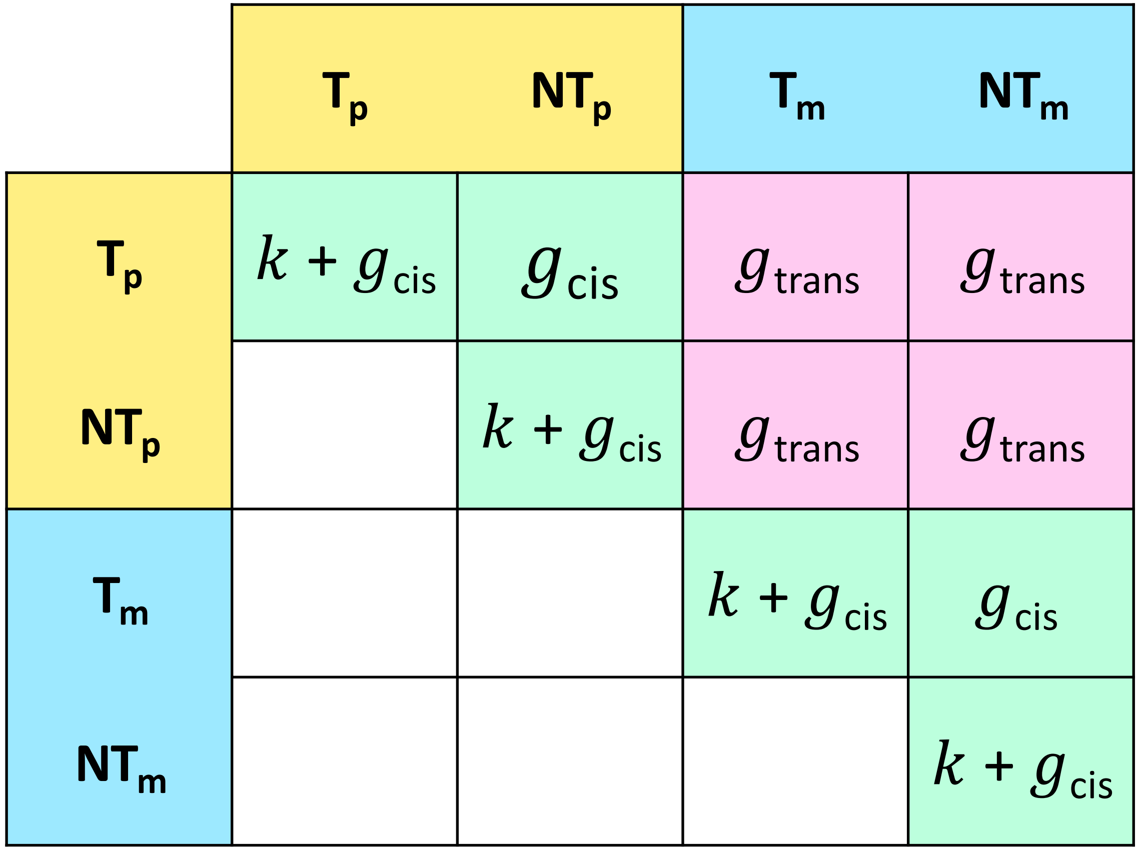
Equilibrium



	T_p	NT_p	T_m	NT_m
T_p	$k + g$	g	g	g
NT_p		$k + g$	g	g
T_m			$k + g$	g
NT_m				$k + g$

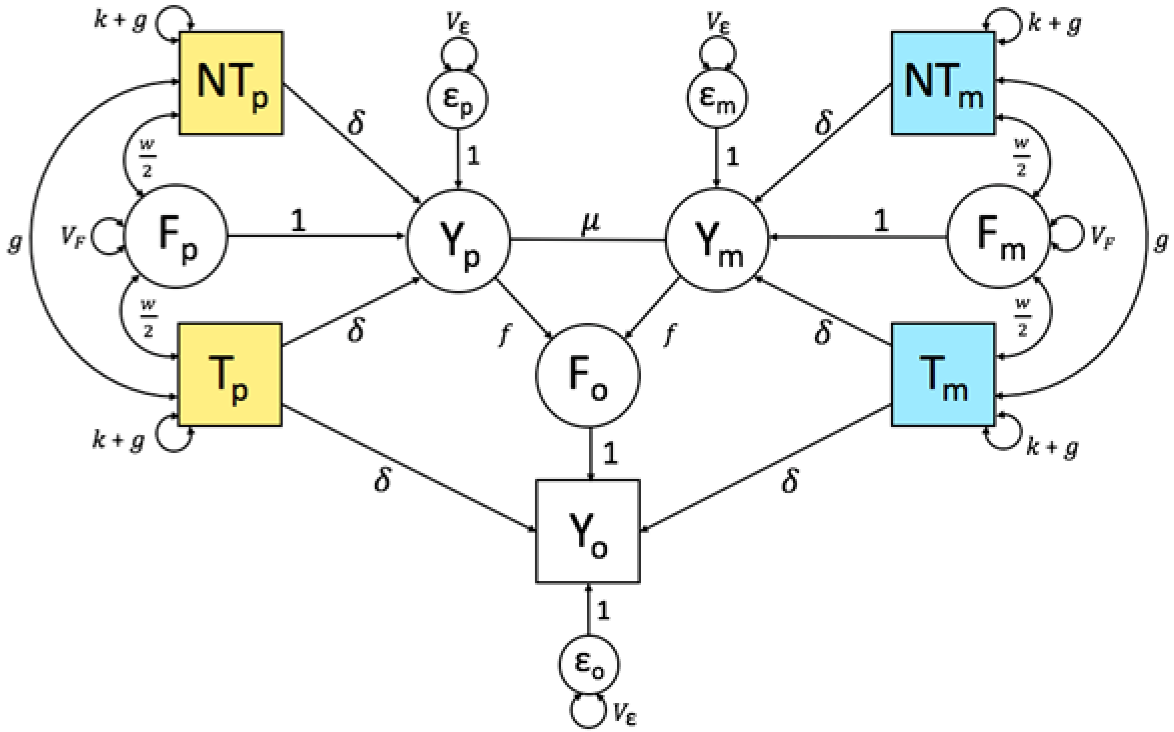
Disequilibrium

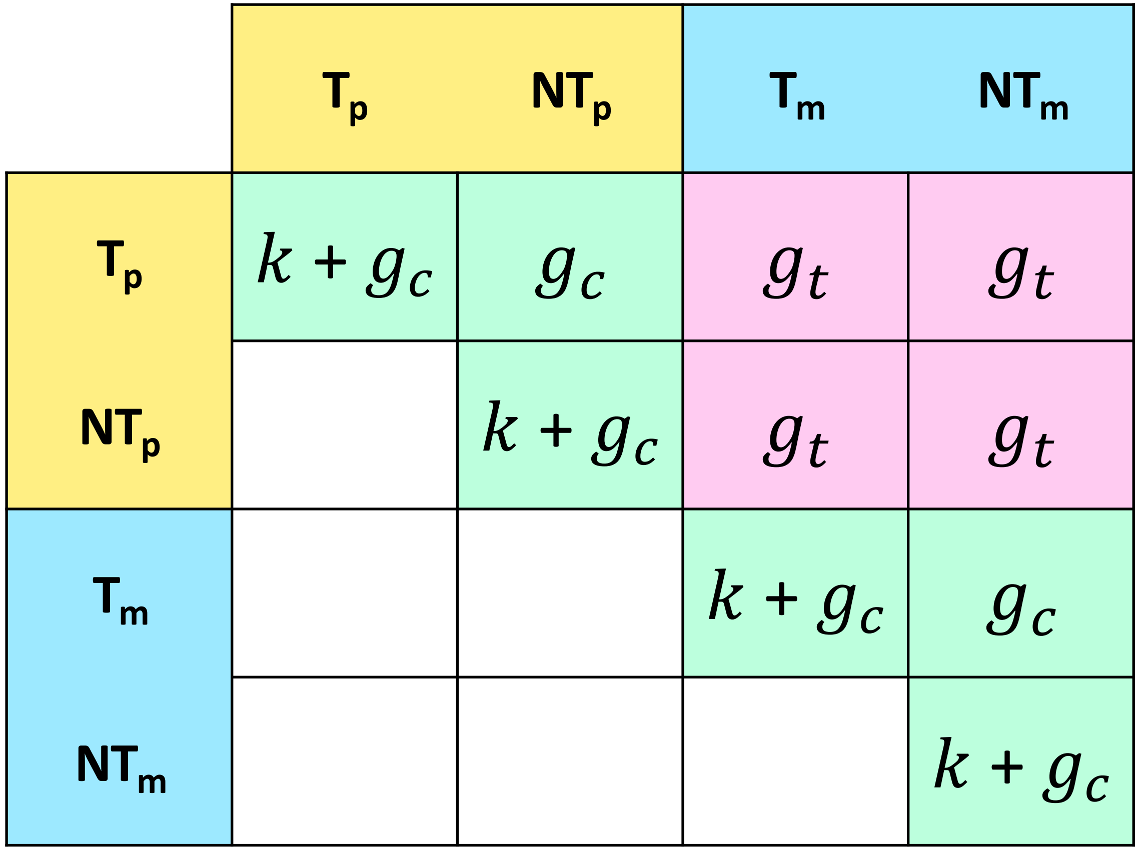




	T_p	NT_p	T_m	NT_m
T_p	$k + g_{cis}$	g_{cis}	g_{trans}	g_{trans}
NT_p		$k + g_{cis}$	g_{trans}	g_{trans}
T_m			$k + g_{cis}$	g_{cis}
NT_m				$k + g_{cis}$

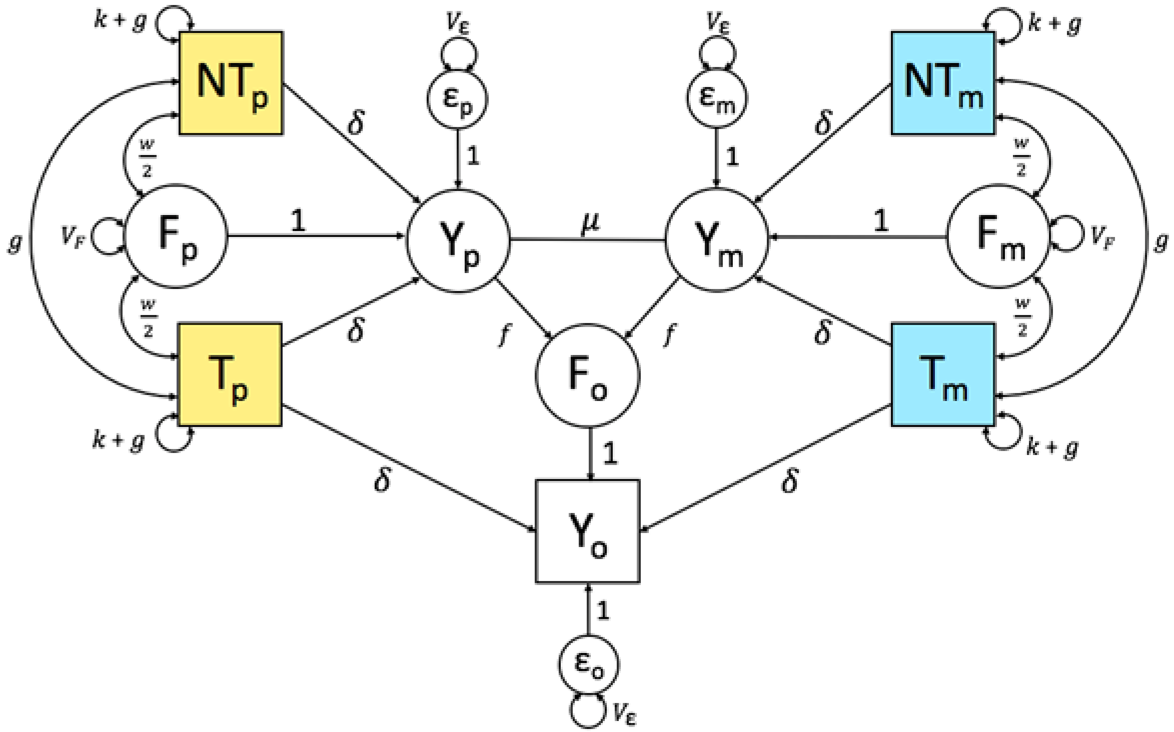
Disequilibrium





	T_p	NT_p	T_m	NT_m
T_p	$k + g_c$	g_c	g_t	g_t
NT_p		$k + g_c$	g_t	g_t
T_m			$k + g_c$	g_c
NT_m				$k + g_c$

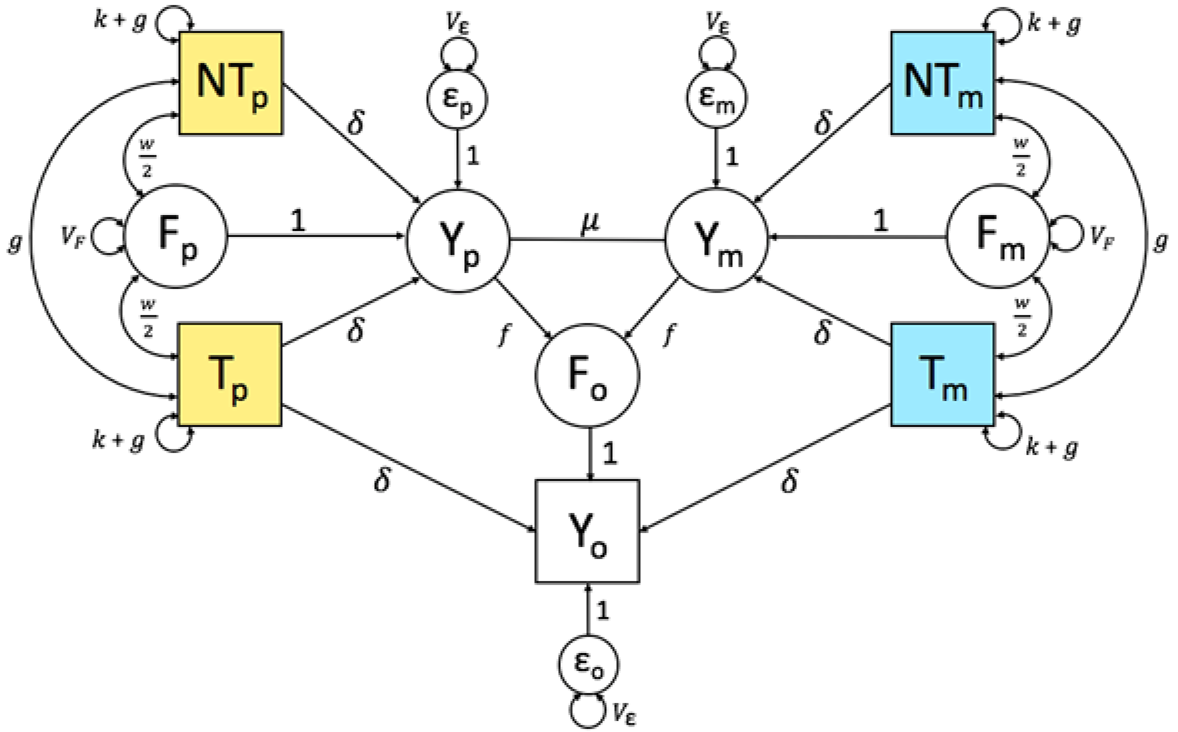
Disequilibrium



	T_p	NT_p	T_m	NT_m
T_p	$k+g_c$	g_c	g_t	g_t
NT_p		$k+g_c$	g_t	g_t
T_m			$k+g_c$	g_c
NT_m				$k+g_c$

g_t is a between-spouse estimate, and tells us about AM in the *parental generation*

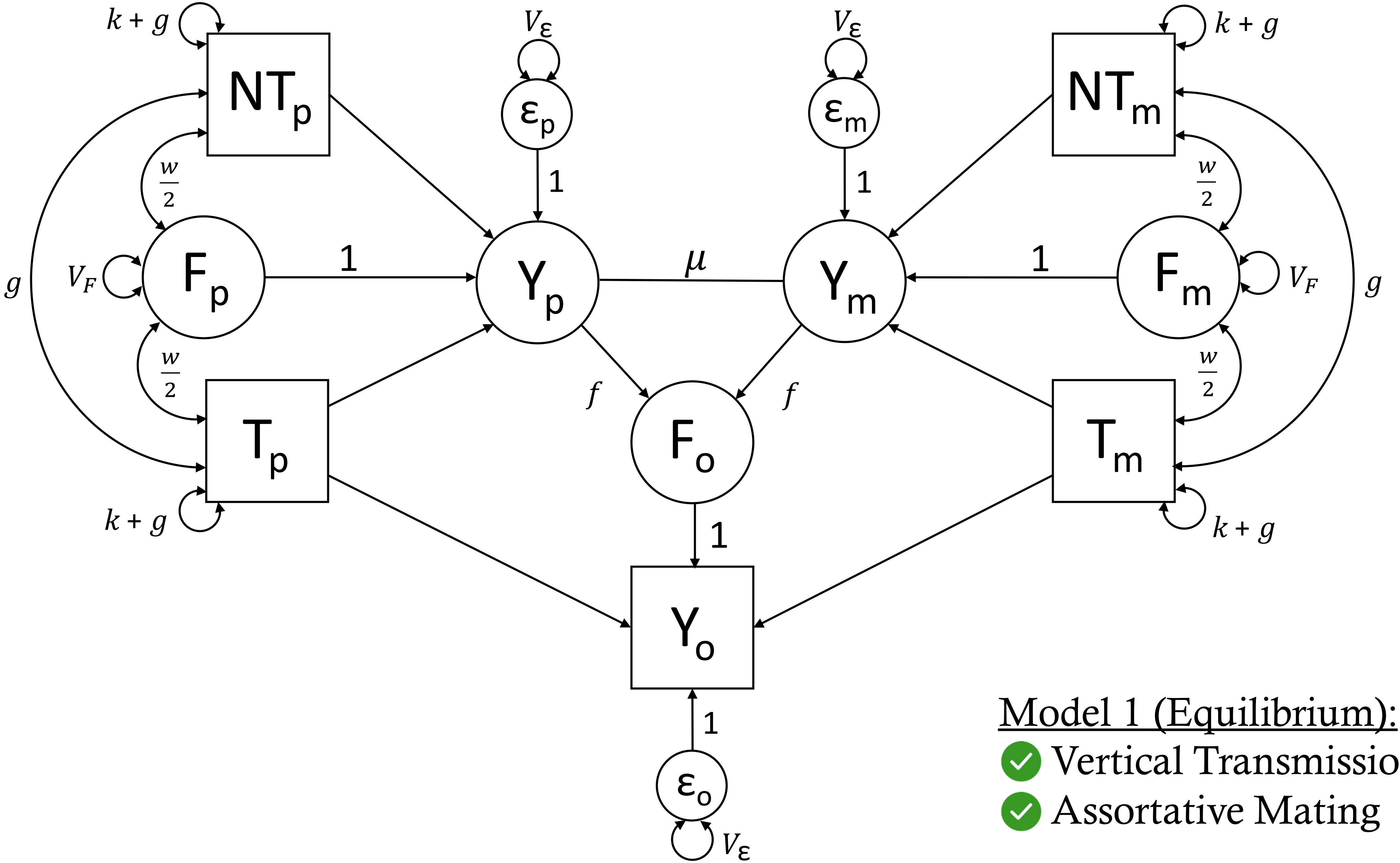
g_c is a within-person estimate, and tells us about AM in the *grandparental generation and earlier*



	T$_p$	**NT$_p$**	**T$_m$**	**NT$_m$**
T$_p$	$k + g_c$	g_c	g_t	g_t
NT$_p$		$k + g_c$	g_t	g_t
T$_m$			$k + g_c$	g_c
NT$_m$				$k + g_c$

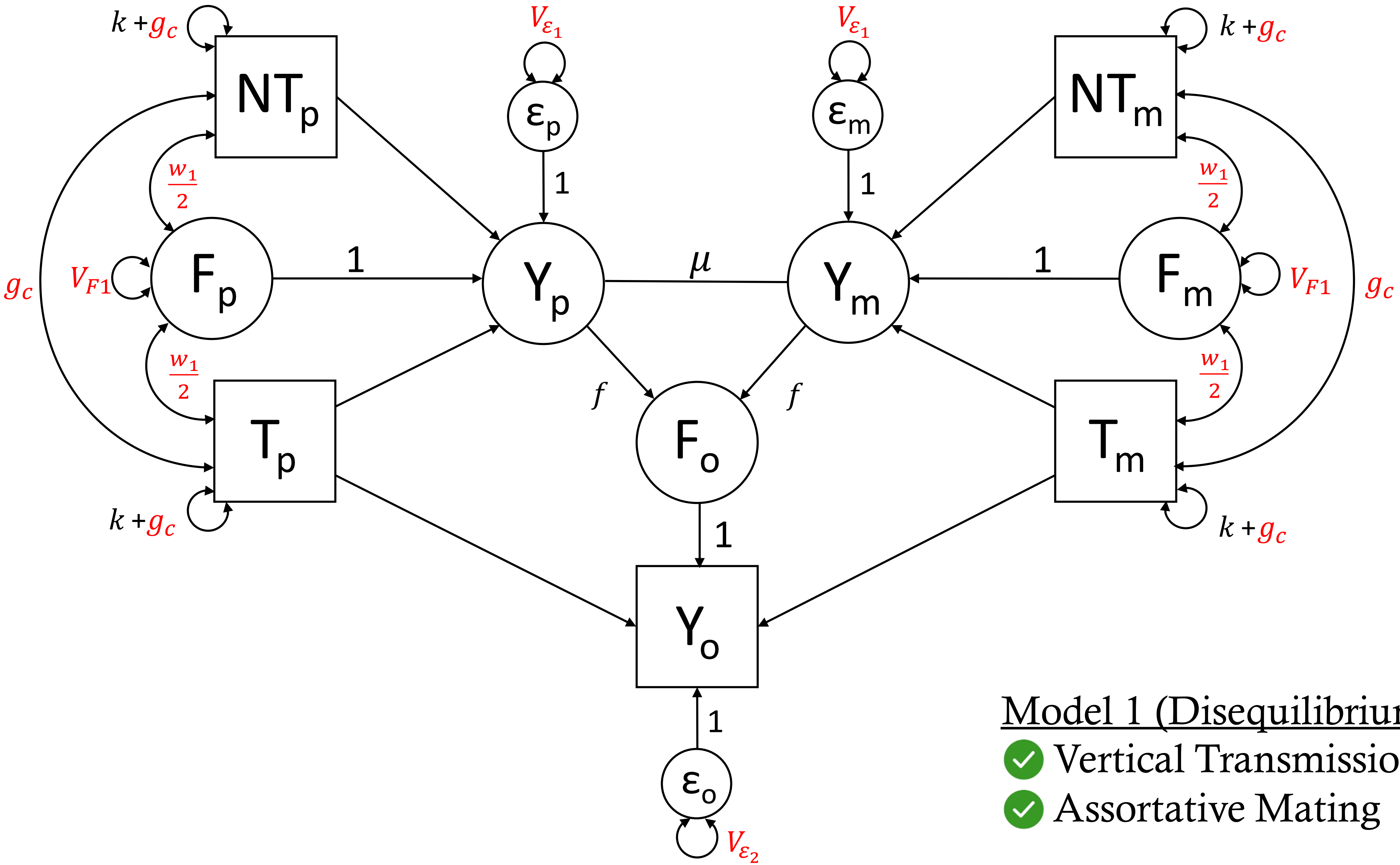
g_t is a between-spouse estimate, and tells us about AM in the *parental generation*

g_c is a within-person estimate, and tells us about AM in the *grandparental generation and earlier*



Model 1 (Equilibrium):

✓ Vertical Transmission
✓ Assortative Mating



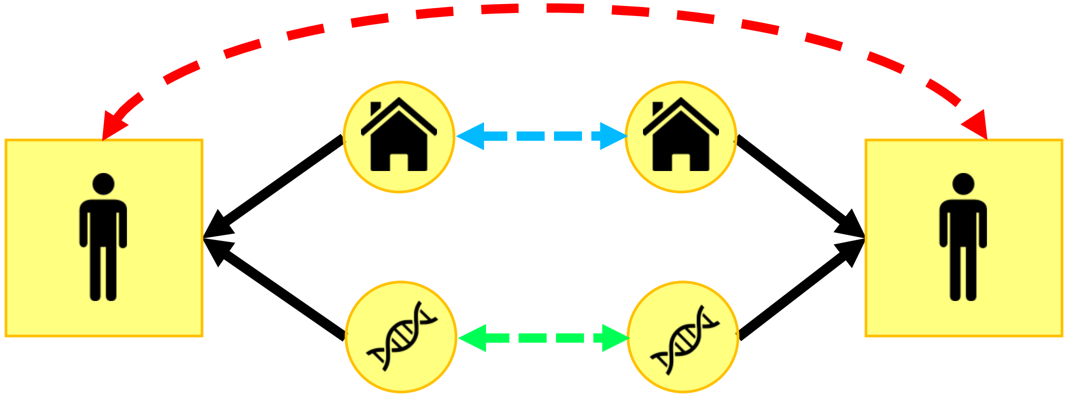
Model 1 (Disequilibrium):

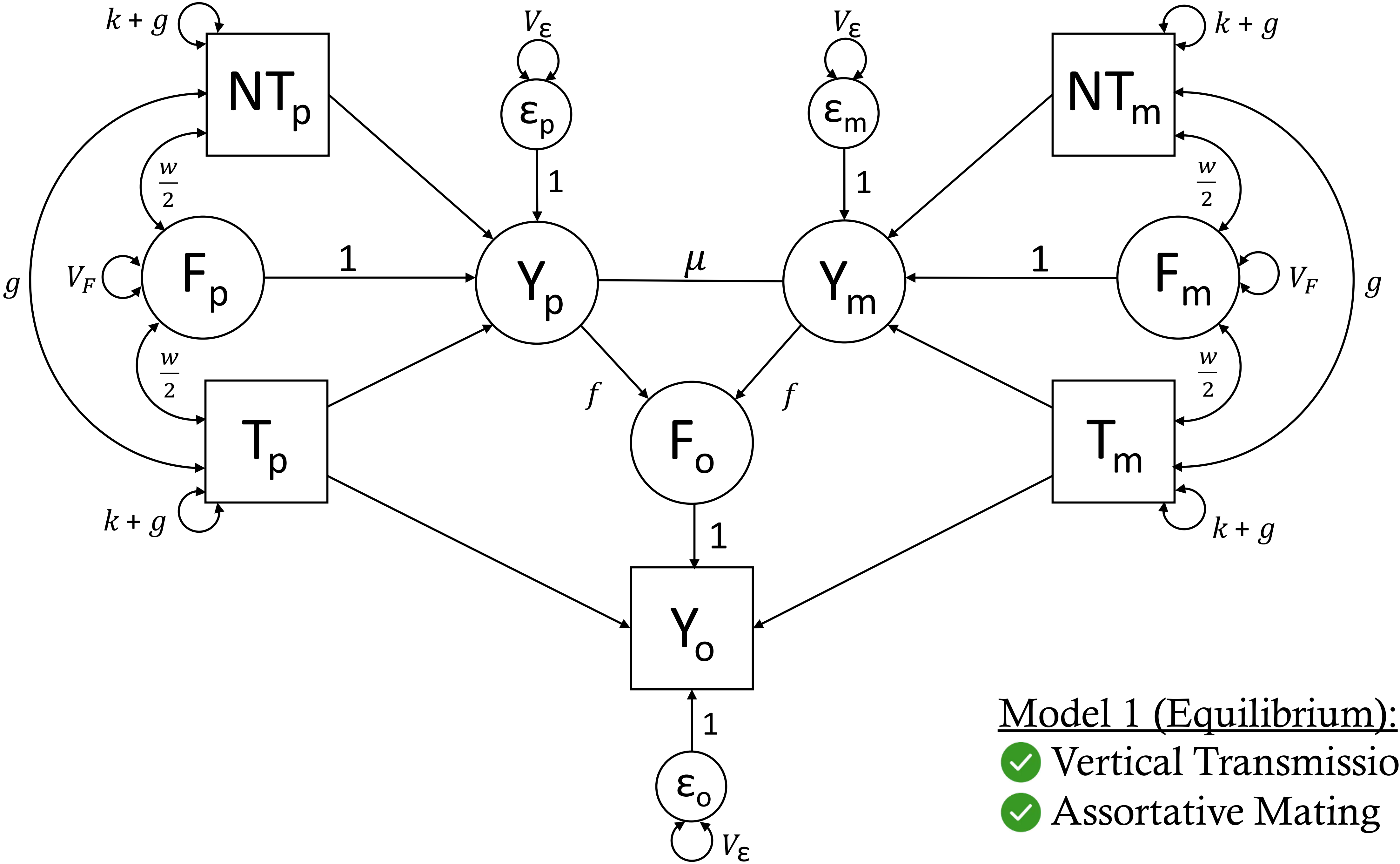
- ✓ Vertical Transmission
- ✓ Assortative Mating

Determining what is causing AM

Types of Assortative Mating

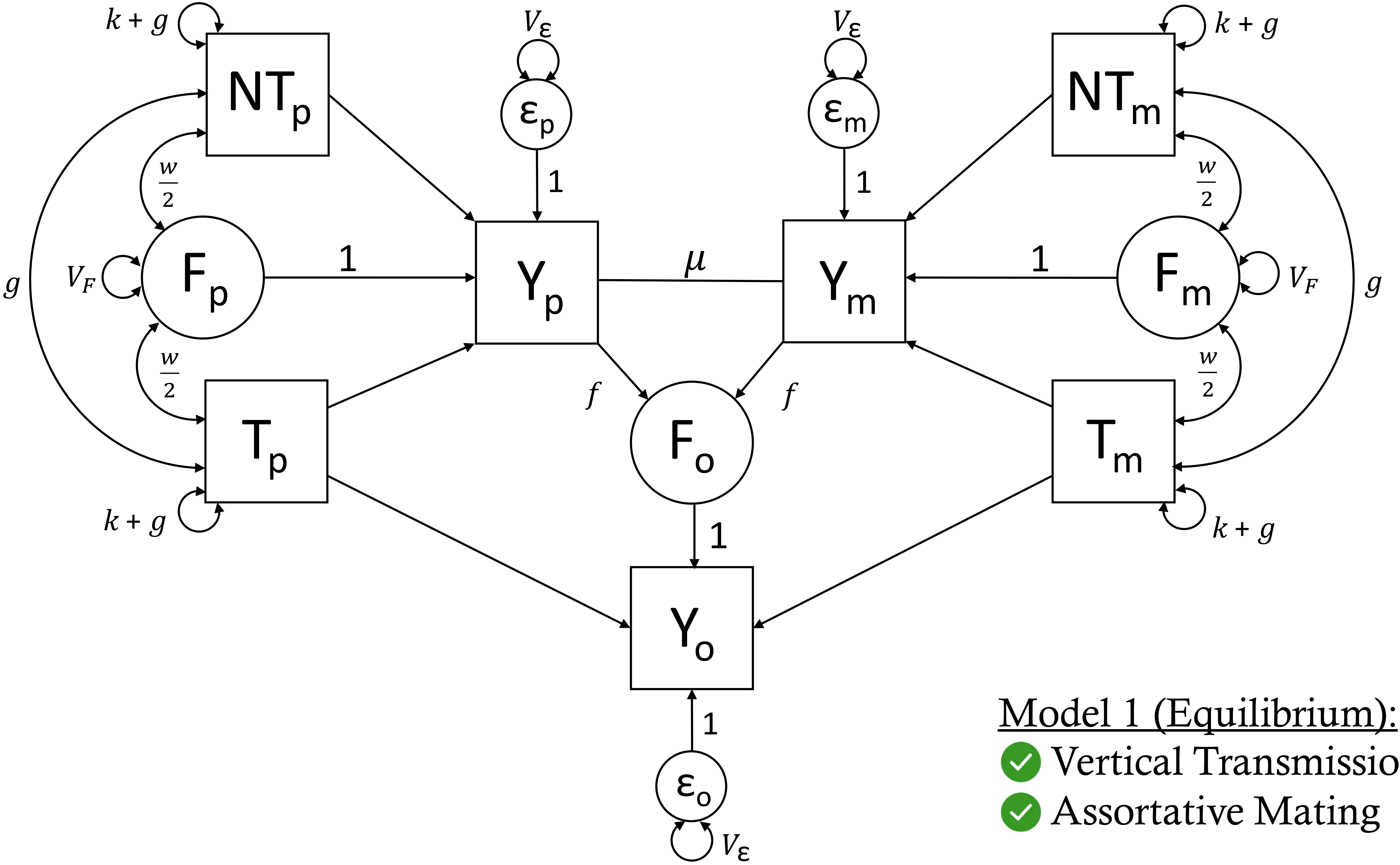
- Phenotypic Homogamy
  - *AKA Primary Phenotypic Assortment*
  - E.g., Mating based on height
- Social Homogamy
  - E.g., Mating based on geographic location
  - Doesn't impact genetic architecture
- Genetic Homogamy
  - Phenotypic homogamy on a different, genetically correlated trait





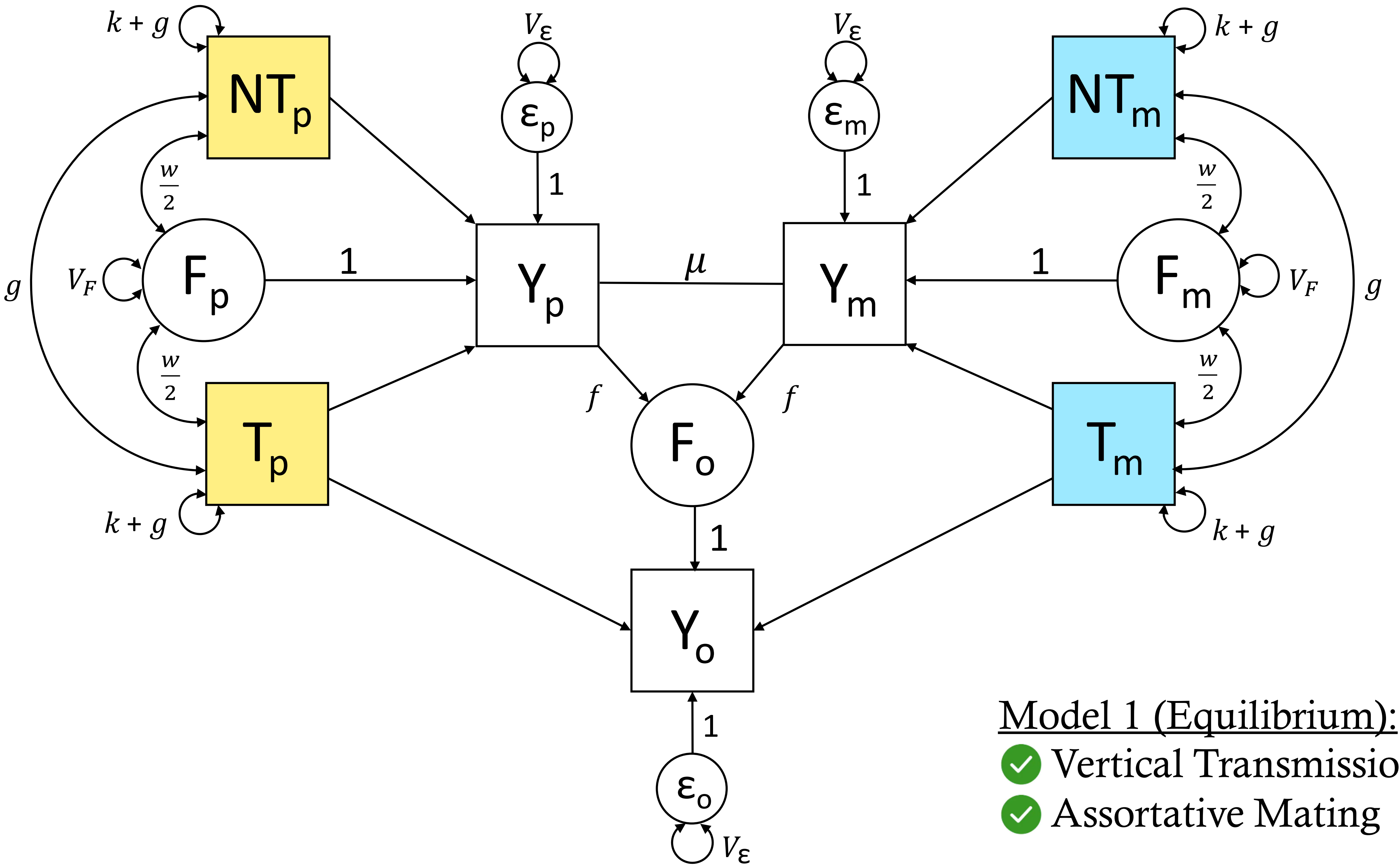
Model 1 (Equilibrium):

- ✓ Vertical Transmission
- ✓ Assortative Mating



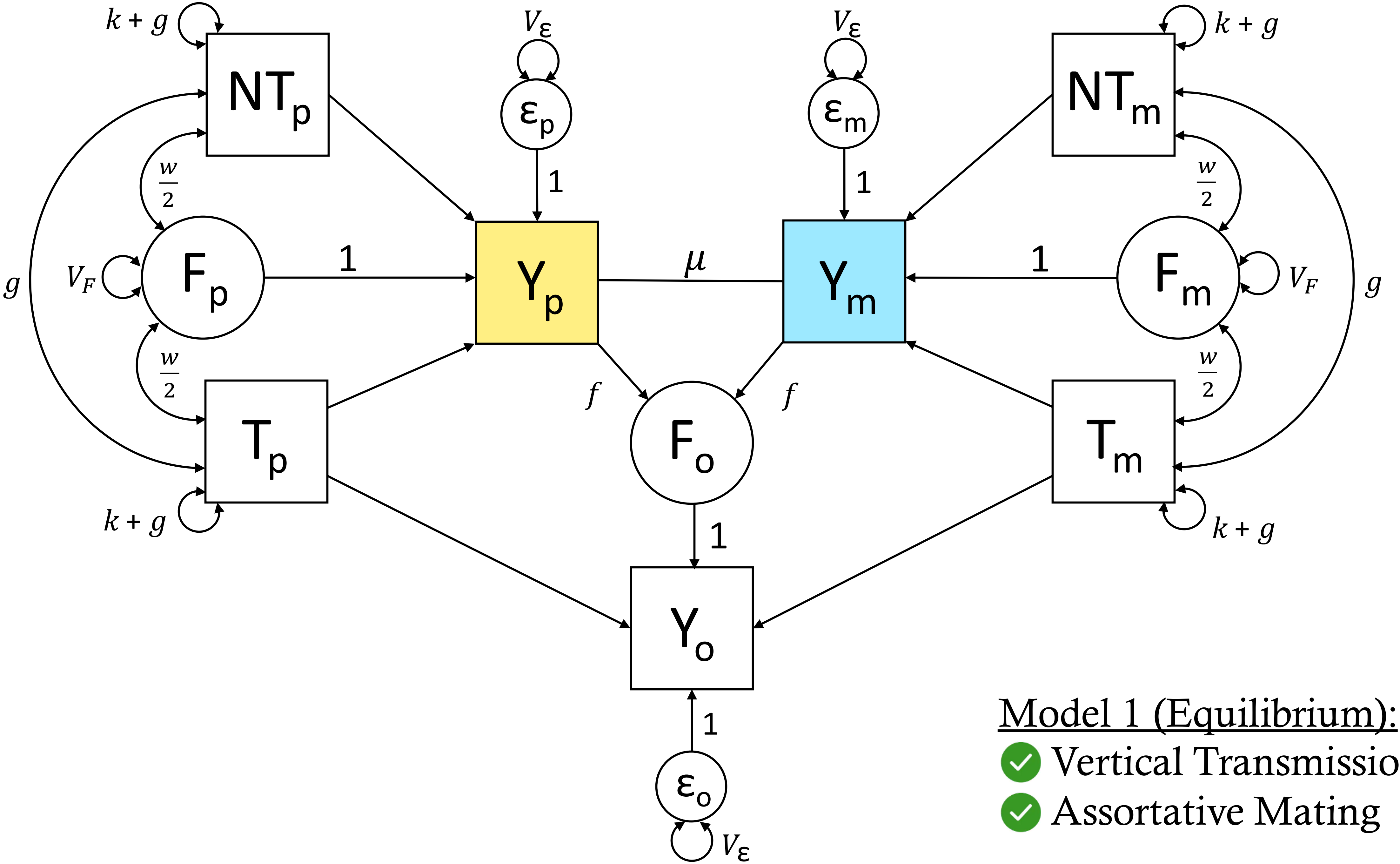
Model 1 (Equilibrium):

- ✓ Vertical Transmission
- ✓ Assortative Mating



<u>Model 1 (Equilibrium):</u>

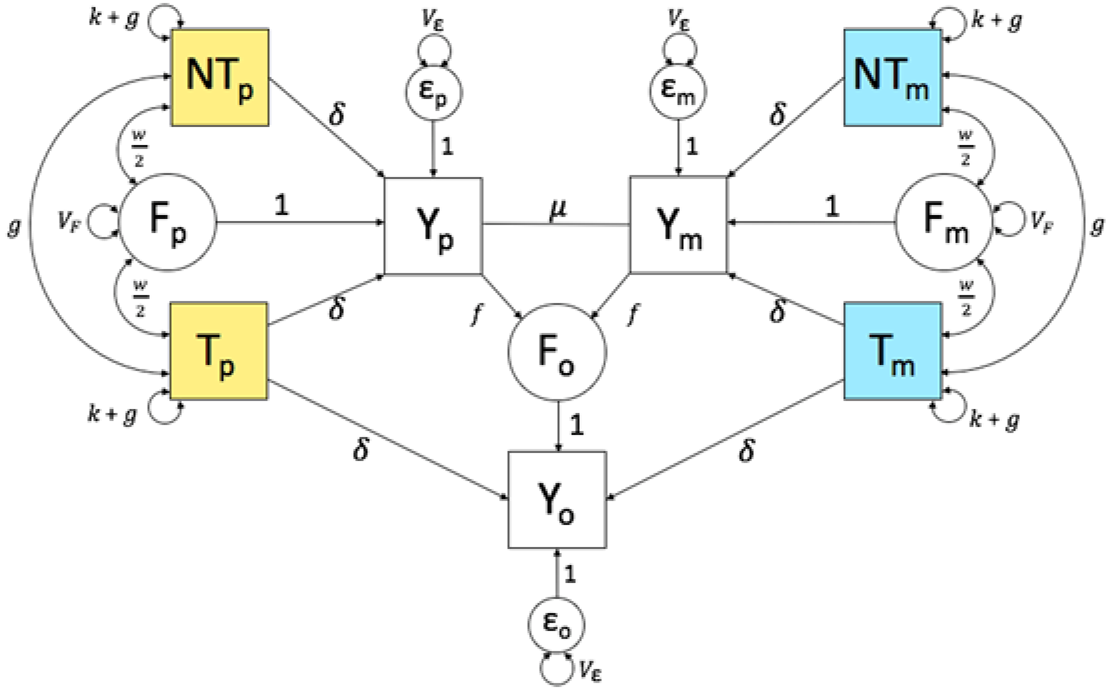
- ✓ Vertical Transmission
- ✓ Assortative Mating



Model 1 (Equilibrium):

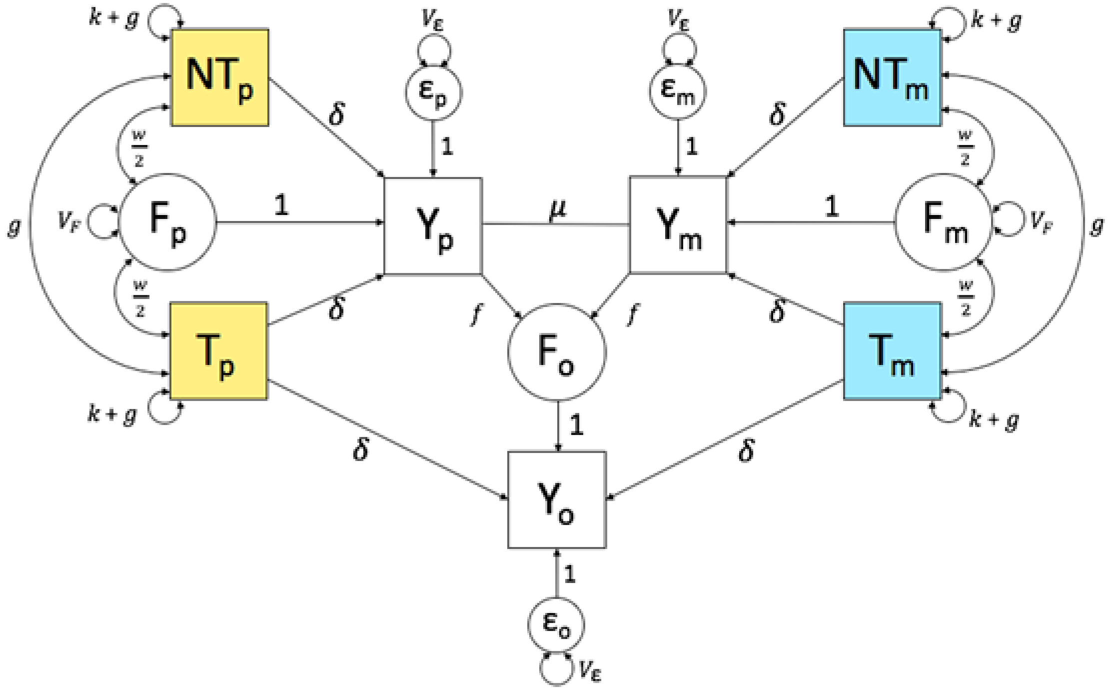
- ✓ Vertical Transmission
- ✓ Assortative Mating

Observed g



	T_p	NT_p	T_m	NT_m
T_p	$k + g$	g	g	g
NT_p		$k + g$	g	g
T_m			$k + g$	g
NT_m				$k + g$

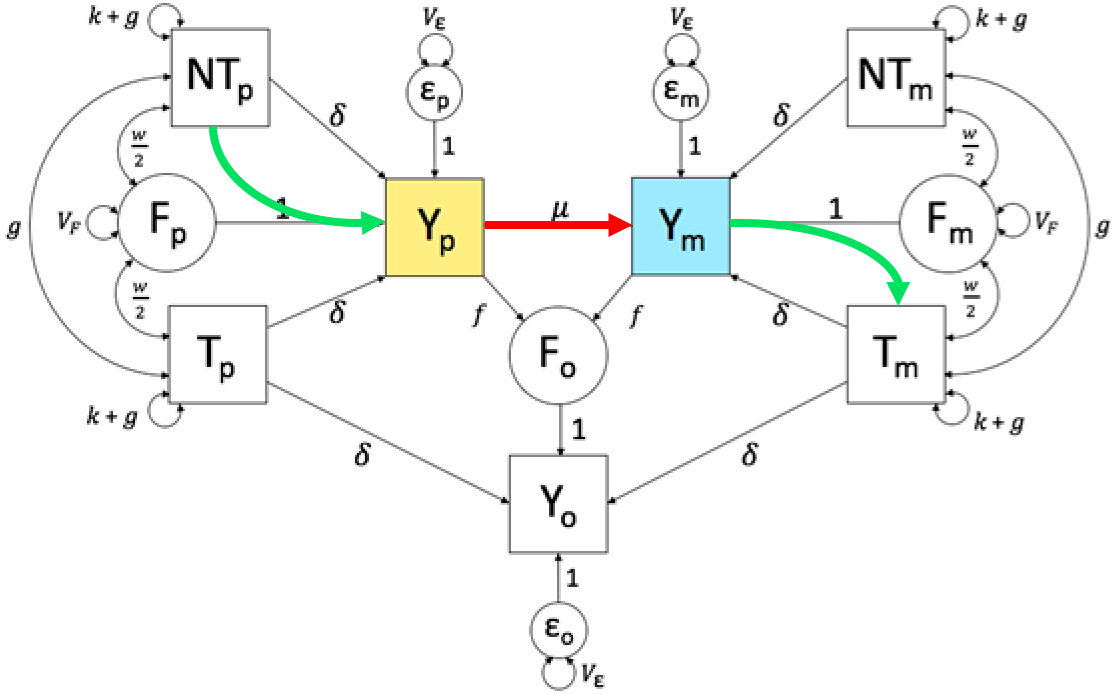
Observed *g*



	T$_p$	NT$_p$	T$_m$	NT$_m$
T$_p$	*k + g*	*g*	*g*	*g*
NT$_p$		*k + g*	*g*	*g*
T$_m$			*k + g*	*g*
NT$_m$				*k + g*

Expected *g*

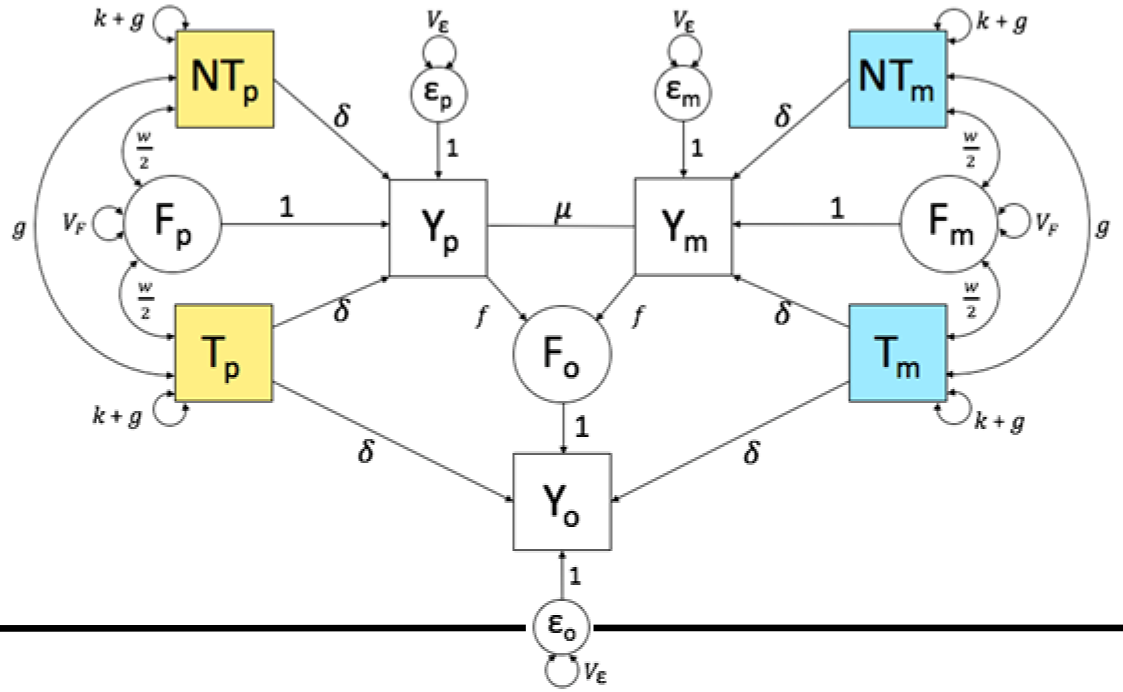
based on spousal phenotypic covariance



$$E(g) = \text{cov}([\text{N}]\text{T}_p\ ,\ [\text{N}]\text{T}_m) = \Omega\mu\Omega = \Omega^2\mu$$

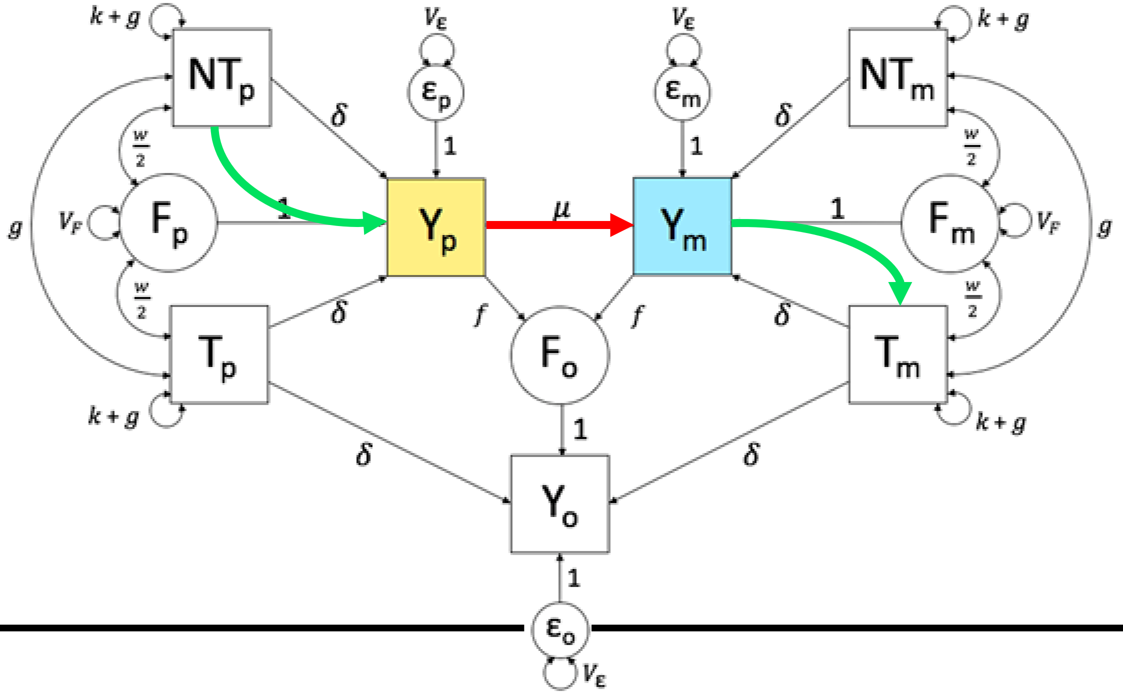
$$\mu = \frac{\text{Spousal Phenotypic Covariance}}{(\text{Spouse Phenotypic Variance})^2}$$

Observed g

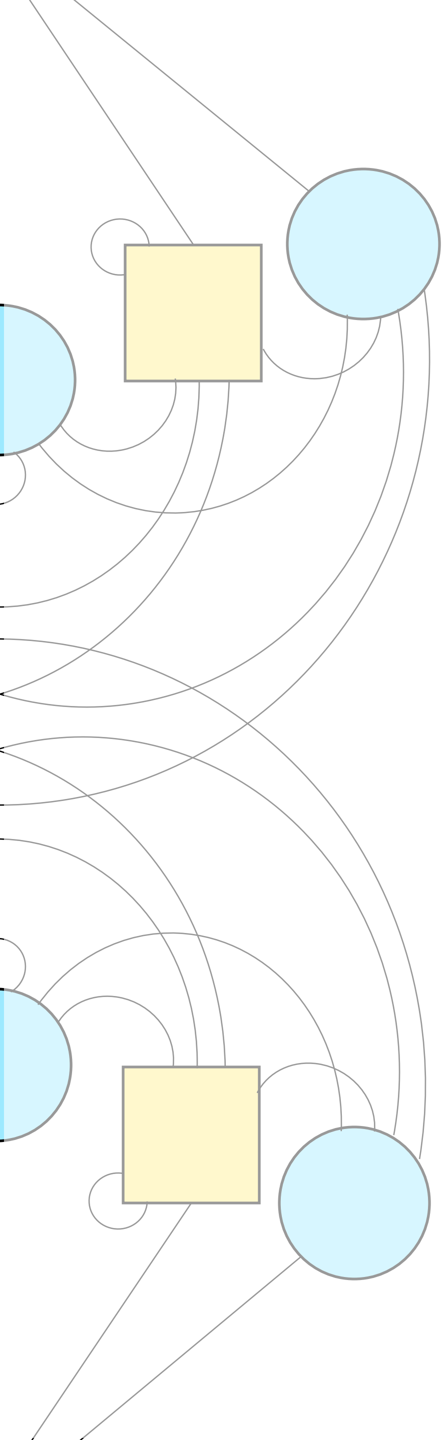


Expected g

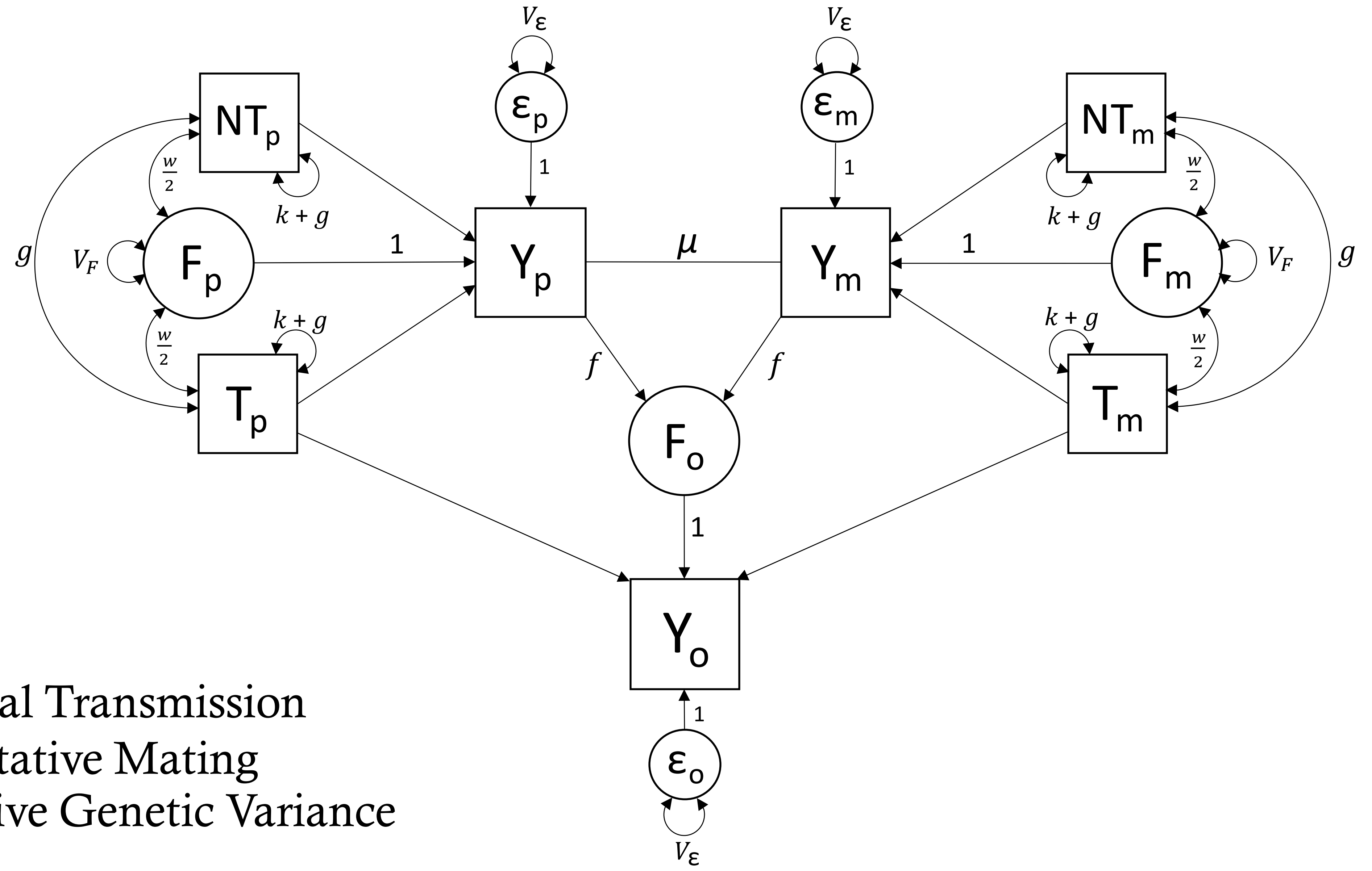
based on spousal phenotypic covariance



- If Observed g = Expected g → Phenotypic Homogamy
- If Observed g < Expected g → Social Homogamy
- If Observed g > Expected g → Genetic Homogamy

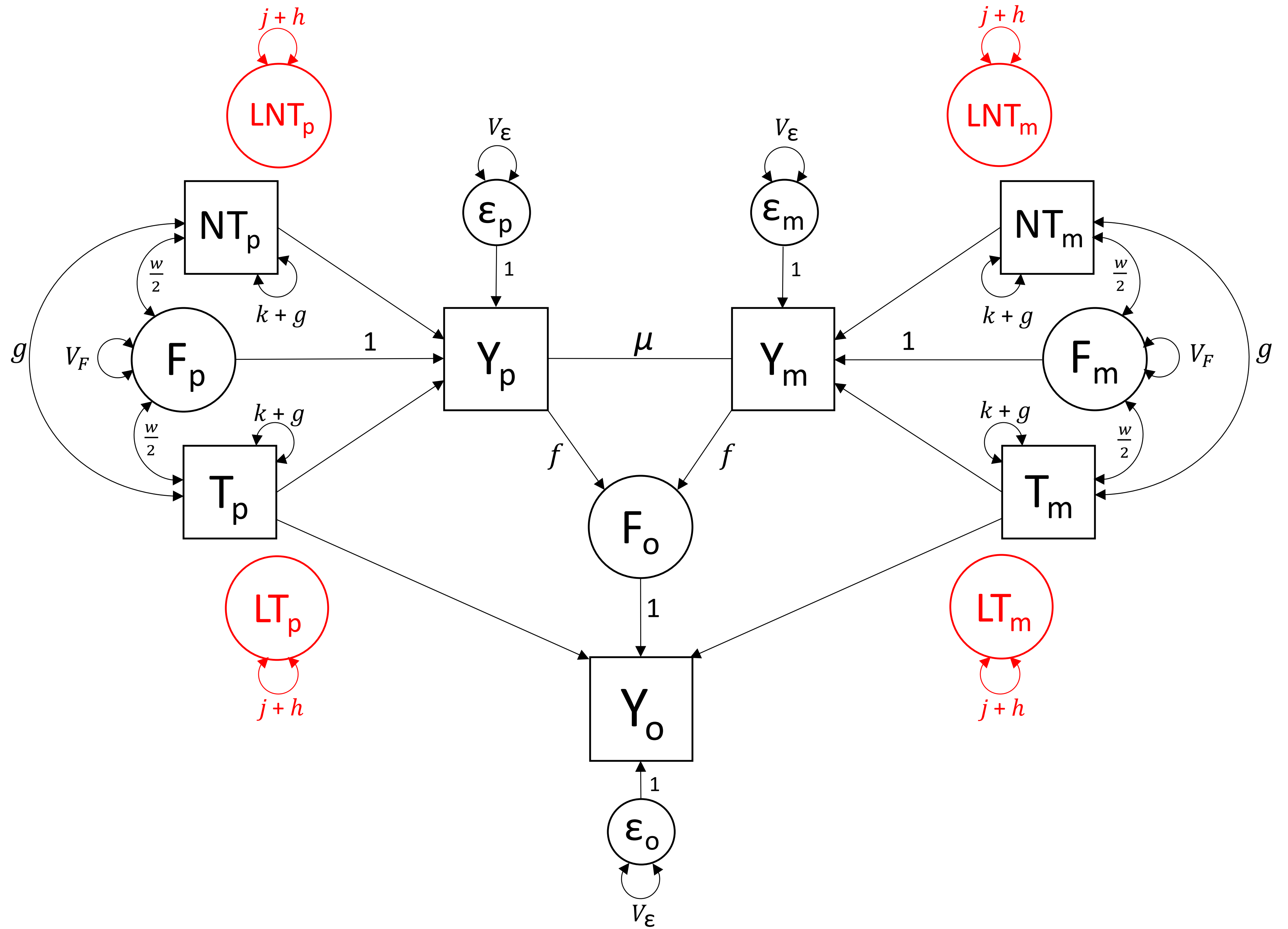


What if the PGS explains a small fraction of total h^2?

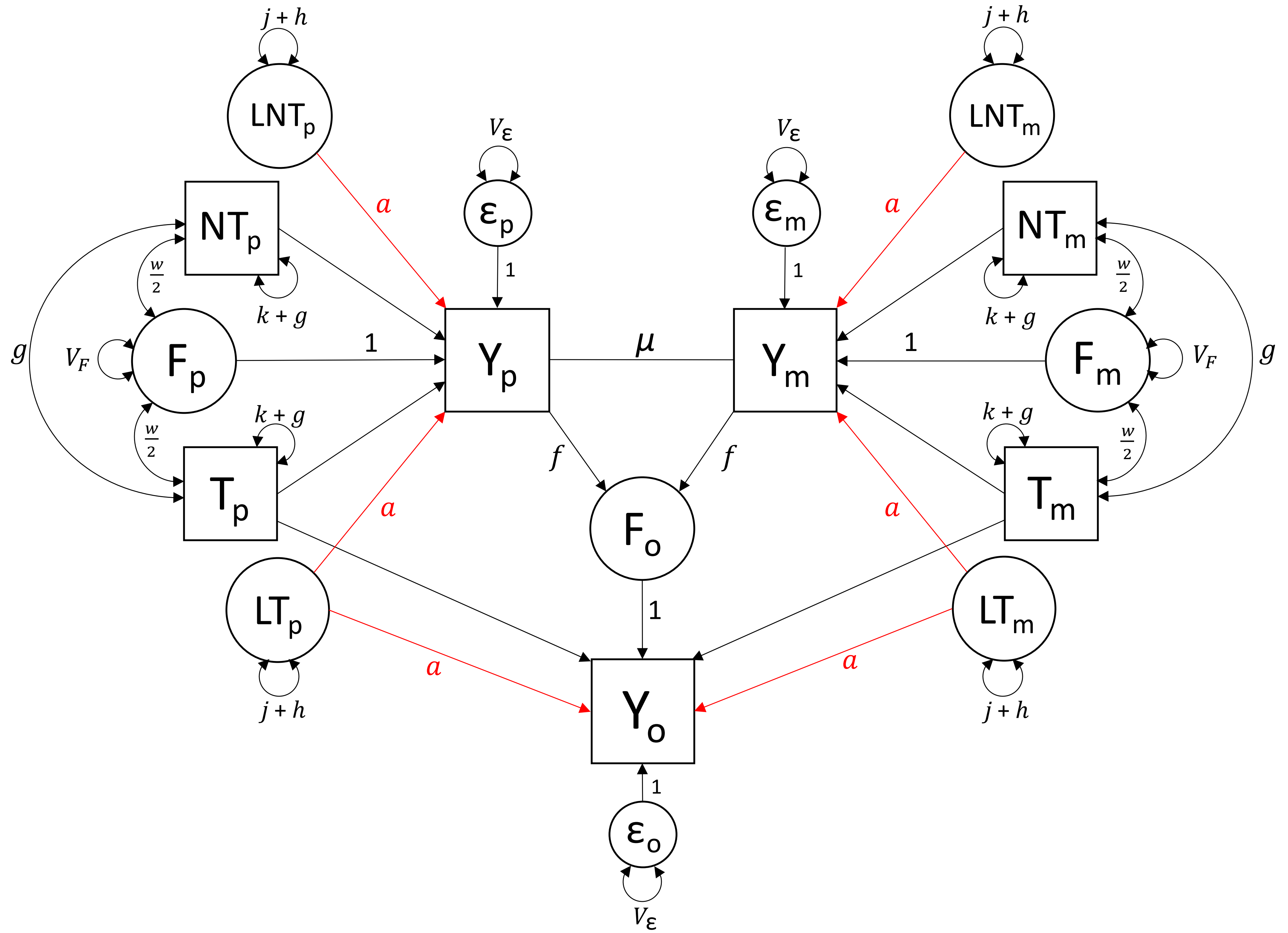


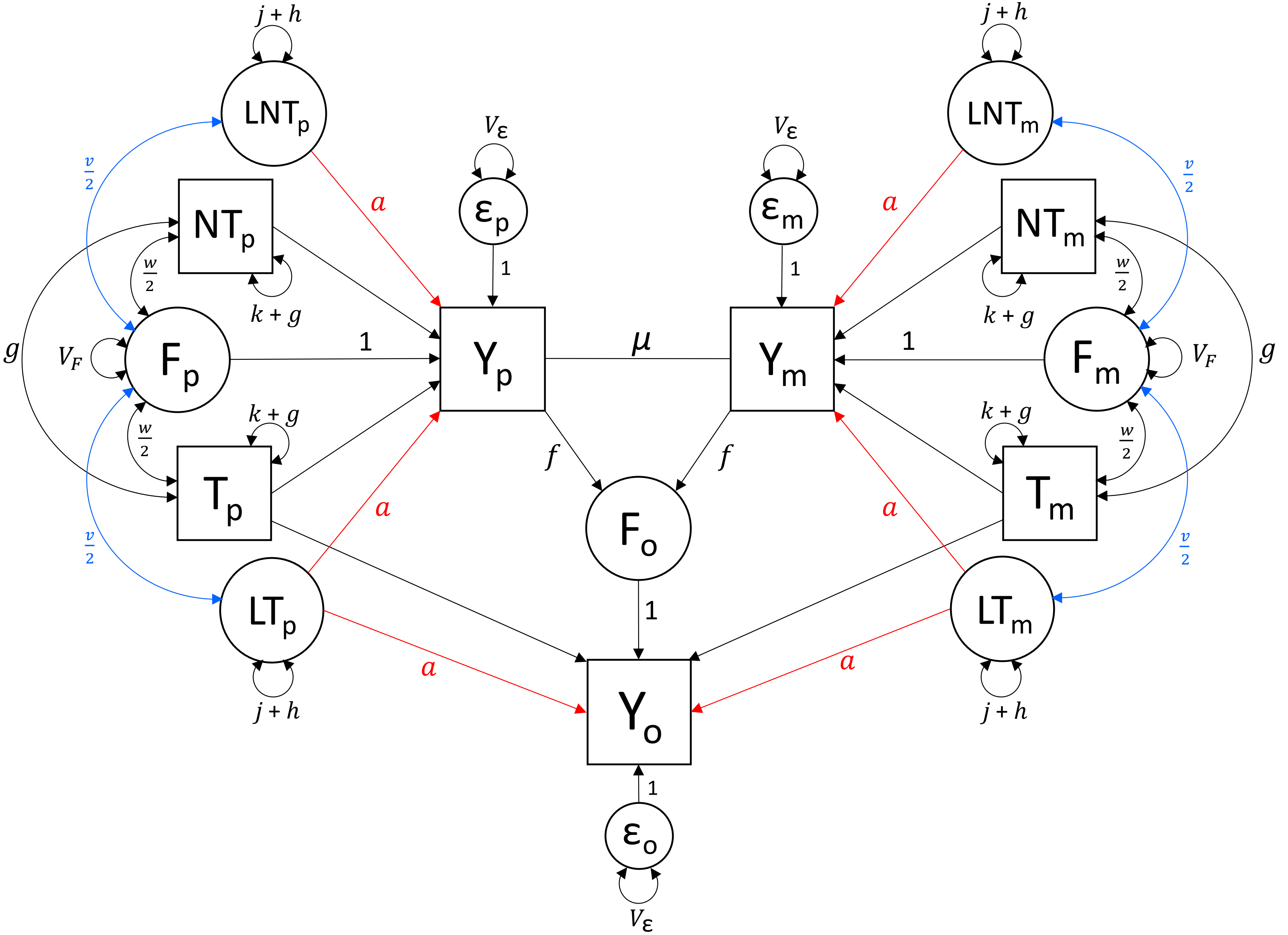
Model 1:

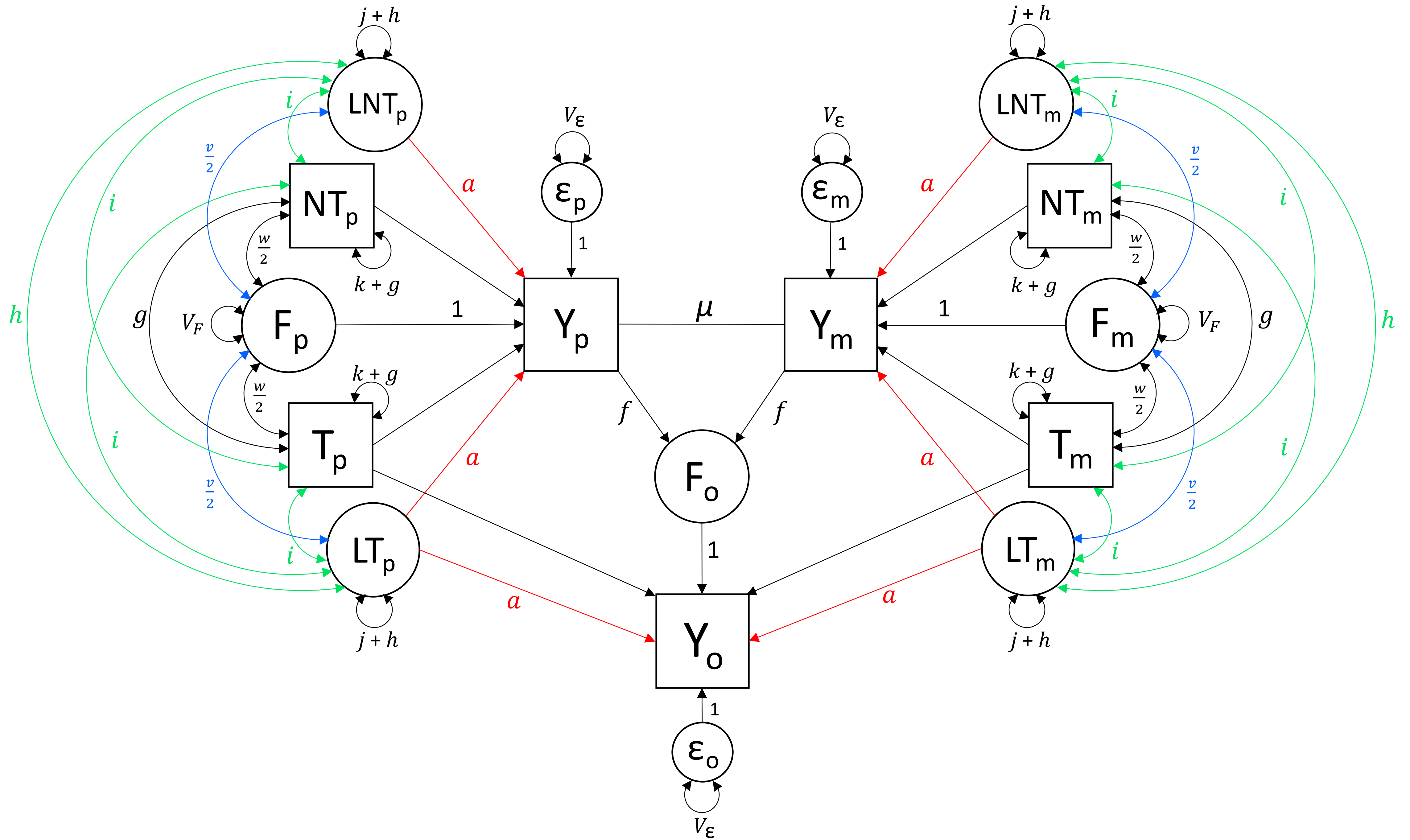
- ✓ Vertical Transmission
- ✓ Assortative Mating
- ✗ Additive Genetic Variance

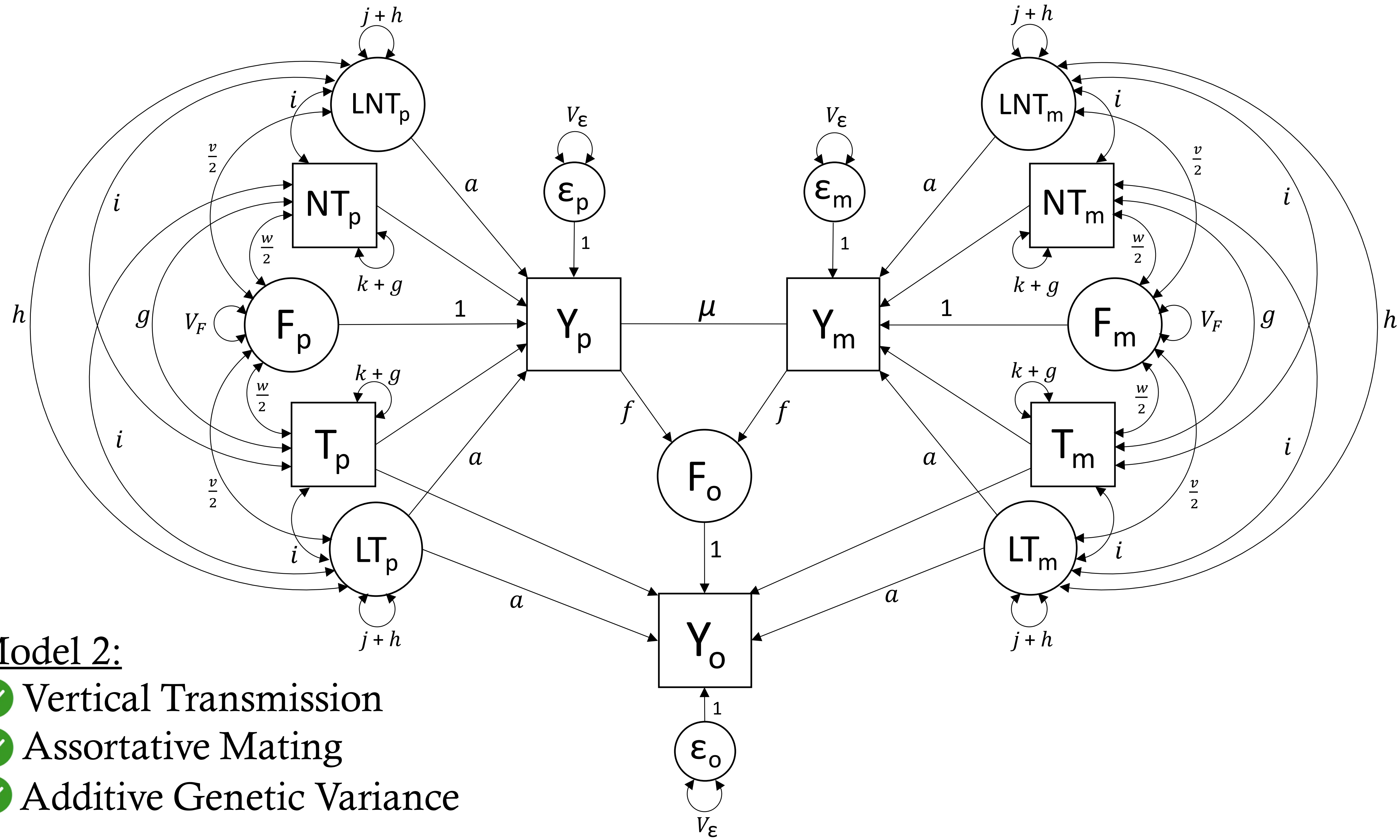


$j + h$
LNT$_p$
$j + h$
LNT$_m$
V_ε
ε_p
V_ε
ε_m
NT$_p$
NT$_m$
$\frac{w}{2}$
$k + g$
1
1
$k + g$
$\frac{w}{2}$
g
V_F
F$_p$
1
Y$_p$
μ
Y$_m$
1
F$_m$
V_F
g
$\frac{w}{2}$
$k + g$
$k + g$
$\frac{w}{2}$
T$_p$
f
f
T$_m$
F$_o$
1
LT$_p$
LT$_m$
$j + h$
$j + h$
Y$_o$
1
ε_o
V_ε


$j + h$
$\mathrm{LNT_p}$
$j + h$
$\mathrm{LNT_m}$
V_ε
V_ε
a
ε_p
ε_m
a
$\mathrm{NT_p}$
$\mathrm{NT_m}$
1
1
$\frac{w}{2}$
$\frac{w}{2}$
$k + g$
$k + g$
g
V_F
$\mathrm{F_p}$
1
$\mathrm{Y_p}$
μ
$\mathrm{Y_m}$
1
$\mathrm{F_m}$
V_F
g
$k + g$
$k + g$
$\frac{w}{2}$
$\frac{w}{2}$
f
f
$\mathrm{T_p}$
$\mathrm{T_m}$
a
a
$\mathrm{F_o}$
$\mathrm{LT_p}$
1
$\mathrm{LT_m}$
a
a
$\mathrm{Y_o}$
$j + h$
$j + h$
1
ε_o
V_ε

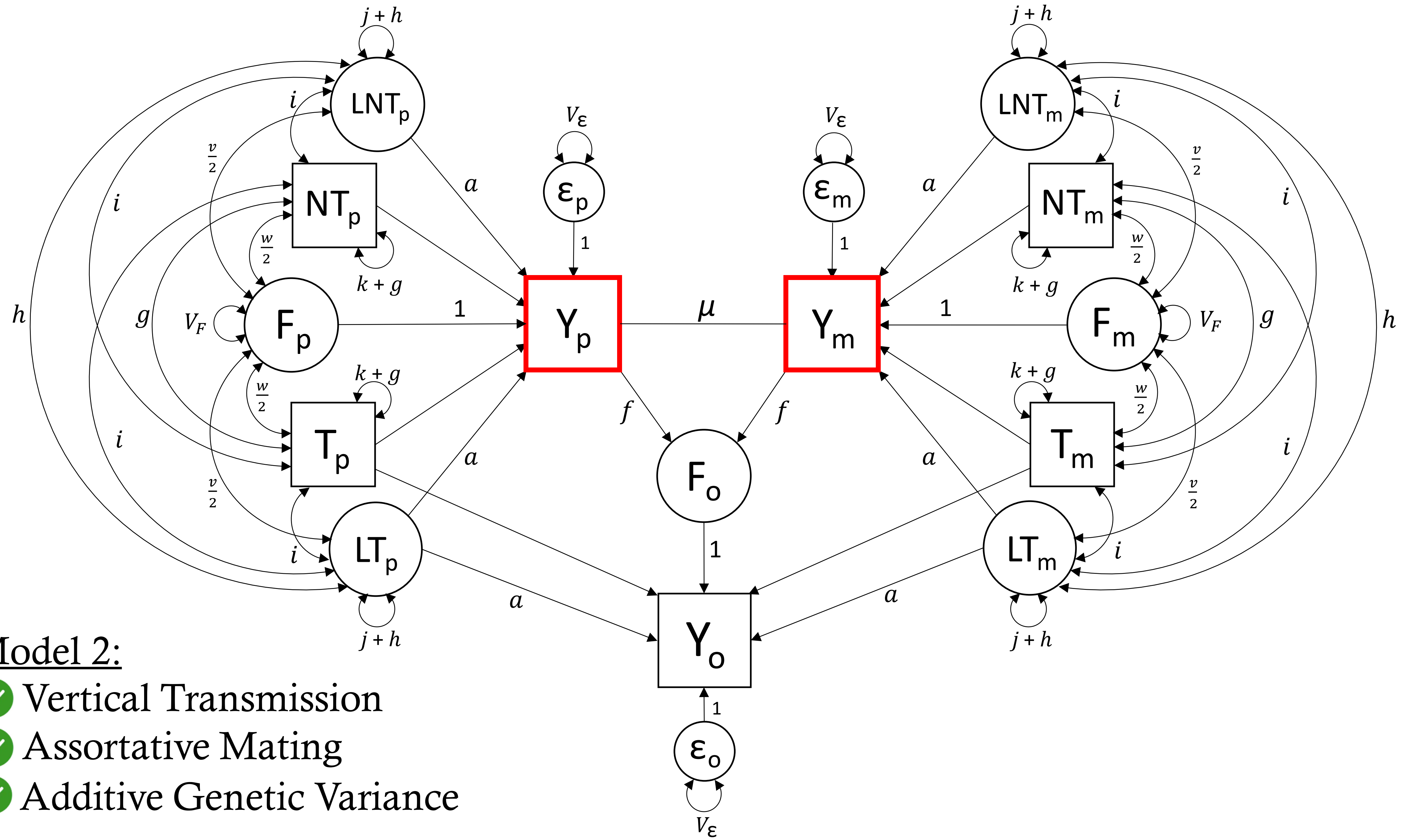

$j + h$
$\mathrm{LNT_p}$
$j + h$
$\mathrm{LNT_m}$
V_ε
ε_p
V_ε
ε_m
$\frac{v}{2}$
a
a
$\frac{v}{2}$
$\mathrm{NT_p}$
1
1
$\mathrm{NT_m}$
$\frac{w}{2}$
$k + g$
$k + g$
$\frac{w}{2}$
g
V_F
$\mathrm{F_p}$
1
$\mathrm{Y_p}$
μ
$\mathrm{Y_m}$
1
$\mathrm{F_m}$
V_F
g
$k + g$
$k + g$
$\frac{w}{2}$
$\frac{w}{2}$
$\mathrm{T_p}$
f
f
$\mathrm{T_m}$
a
$\mathrm{F_o}$
a
$\frac{v}{2}$
$\frac{v}{2}$
$\mathrm{LT_p}$
1
$\mathrm{LT_m}$
a
a
$j + h$
$\mathrm{Y_o}$
$j + h$
1
ε_o
V_ε


j + h
j + h
LNT$_p$
LNT$_m$
V_ε
V_ε
i
i
a
a
ε$_p$
ε$_m$
$\frac{v}{2}$
$\frac{v}{2}$
i
i
NT$_p$
NT$_m$
1
1
$\frac{w}{2}$
$\frac{w}{2}$
k + g
k + g
h
g
V_F
F$_p$
1
Y$_p$
μ
Y$_m$
1
F$_m$
V_F
g
h
k + g
k + g
$\frac{w}{2}$
$\frac{w}{2}$
f
f
i
T$_p$
a
F$_o$
a
T$_m$
i
$\frac{v}{2}$
$\frac{v}{2}$
i
LT$_p$
1
LT$_m$
i
a
a
j + h
Y$_o$
j + h
1
ε$_o$
V_ε



Model 2:

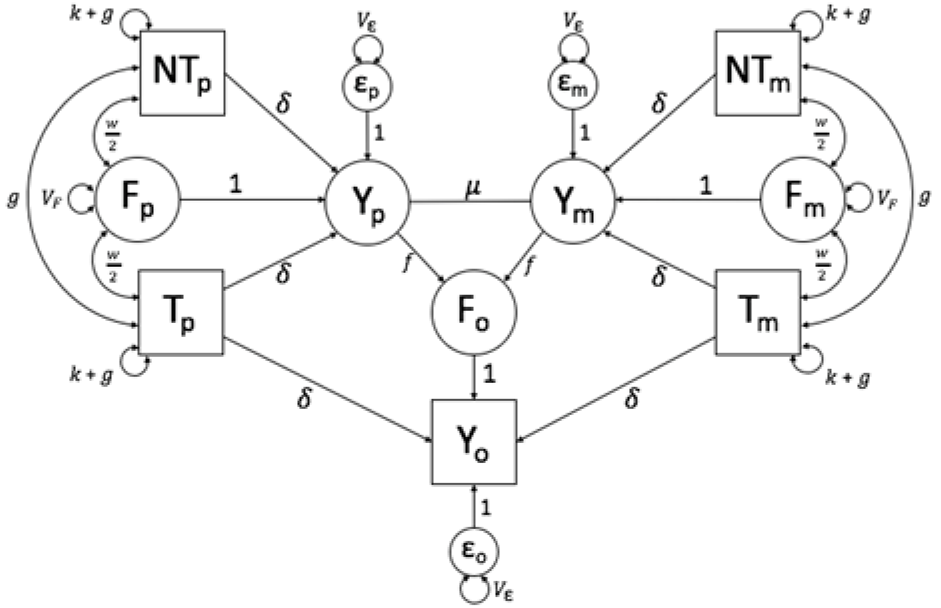
- ✓ Vertical Transmission
- ✓ Assortative Mating
- ✓ Additive Genetic Variance



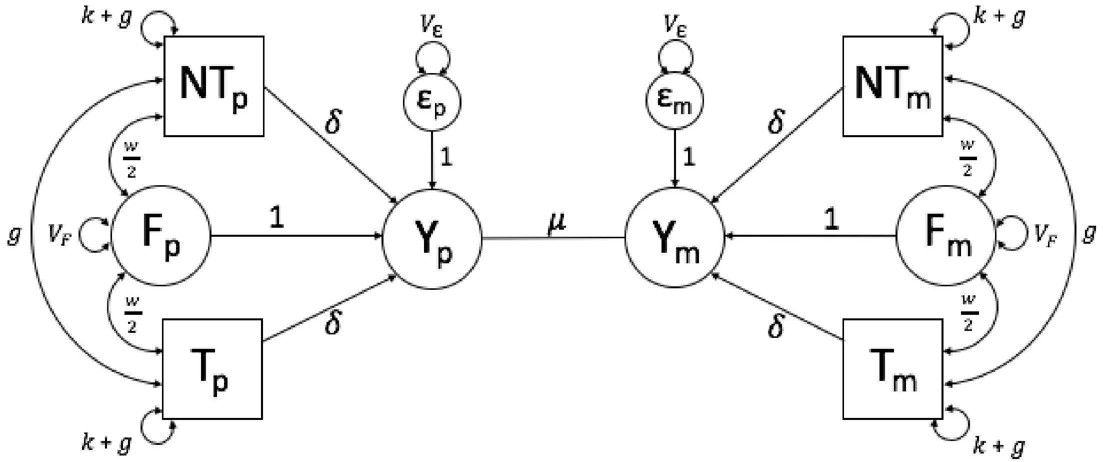
Model 2:

- ✓ Vertical Transmission
- ✓ Assortative Mating
- ✓ Additive Genetic Variance

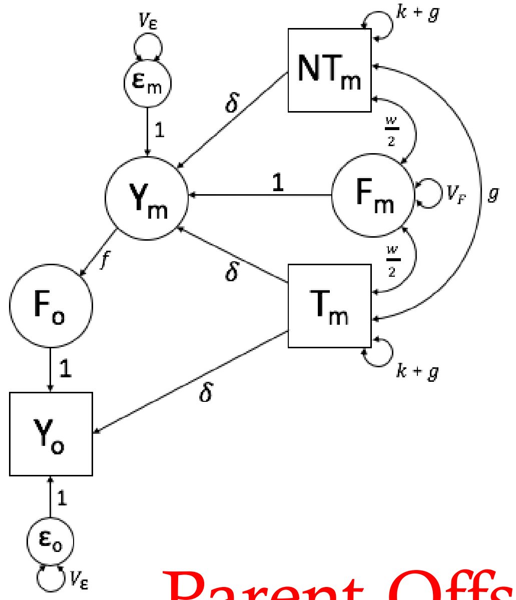
Some other extensions we’re working on



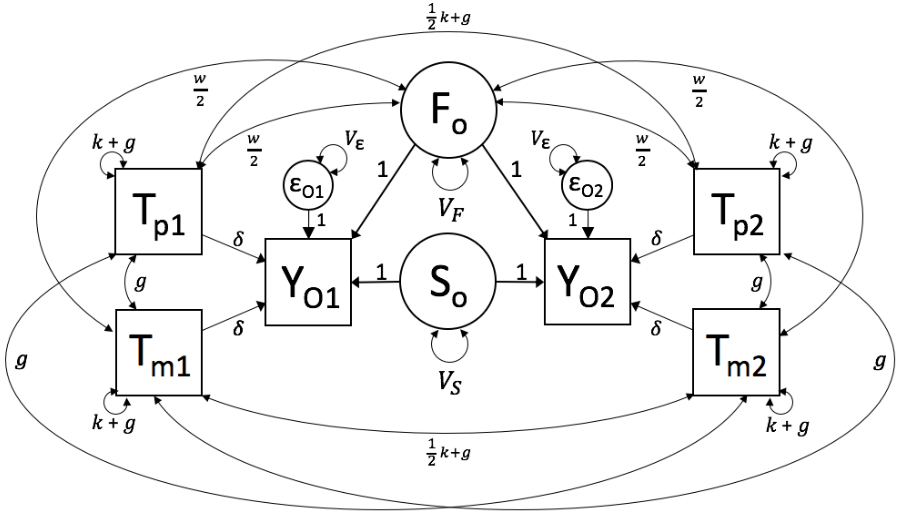
Trio



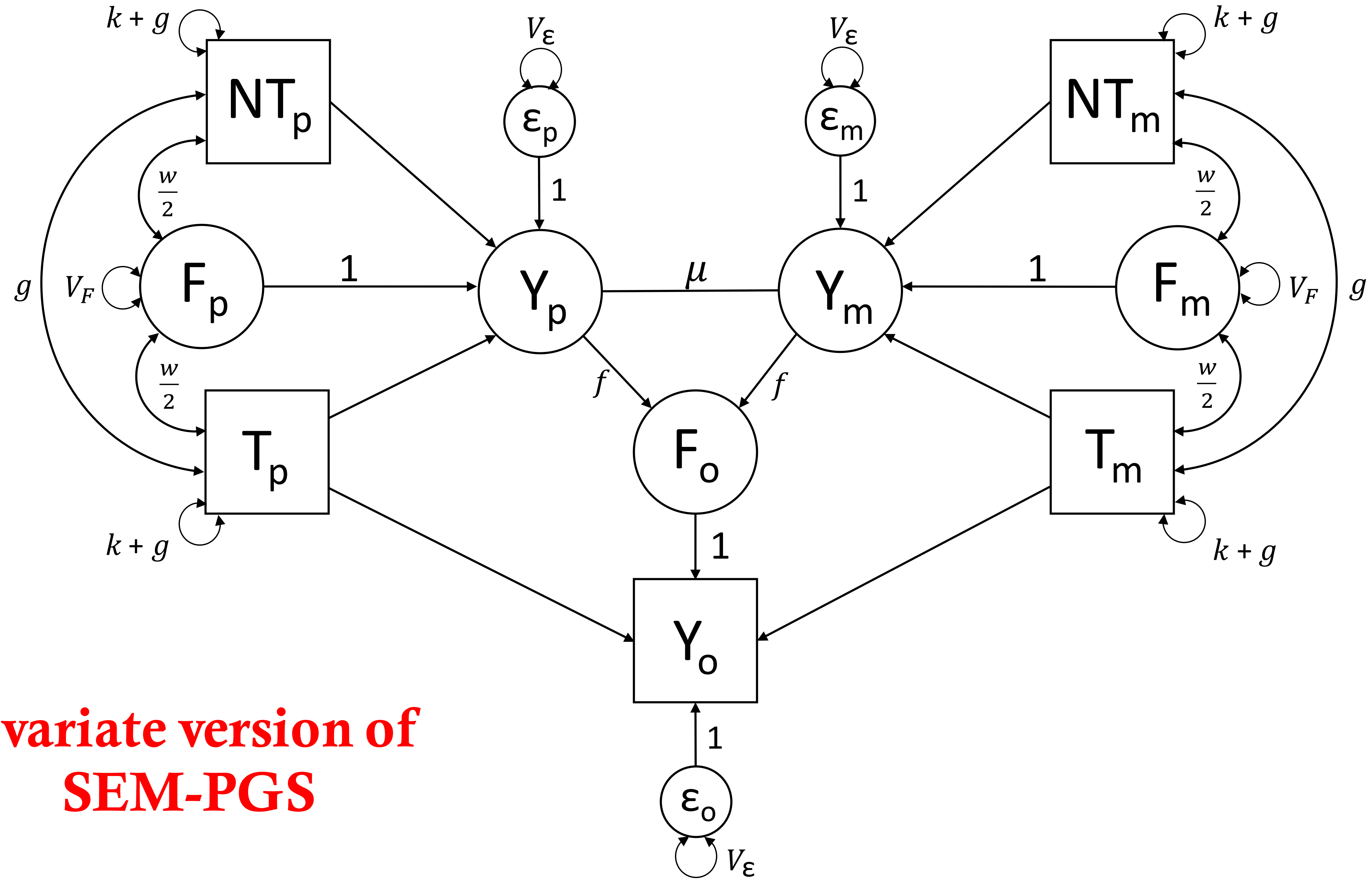
Spouse



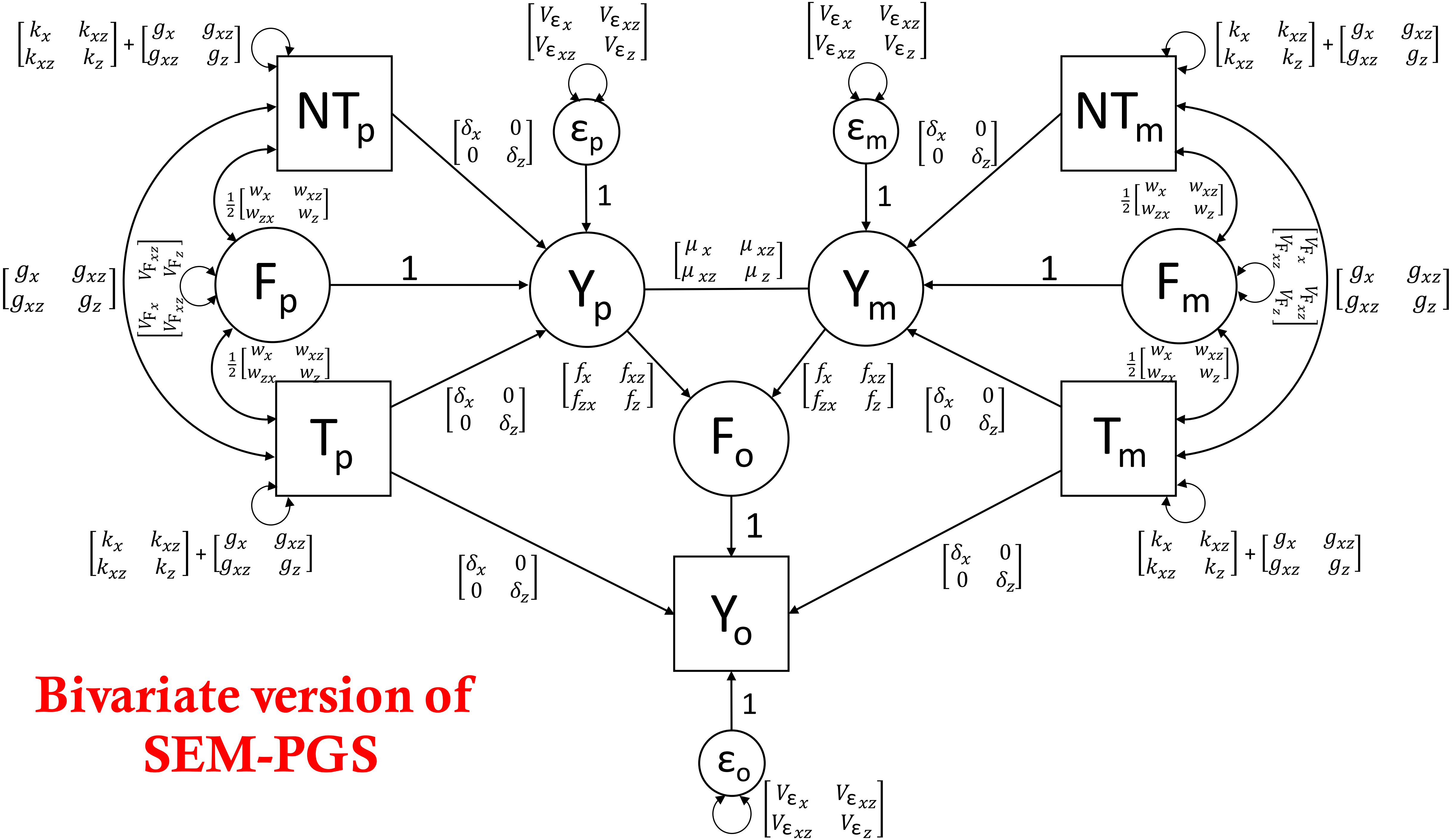
Parent-Offspring



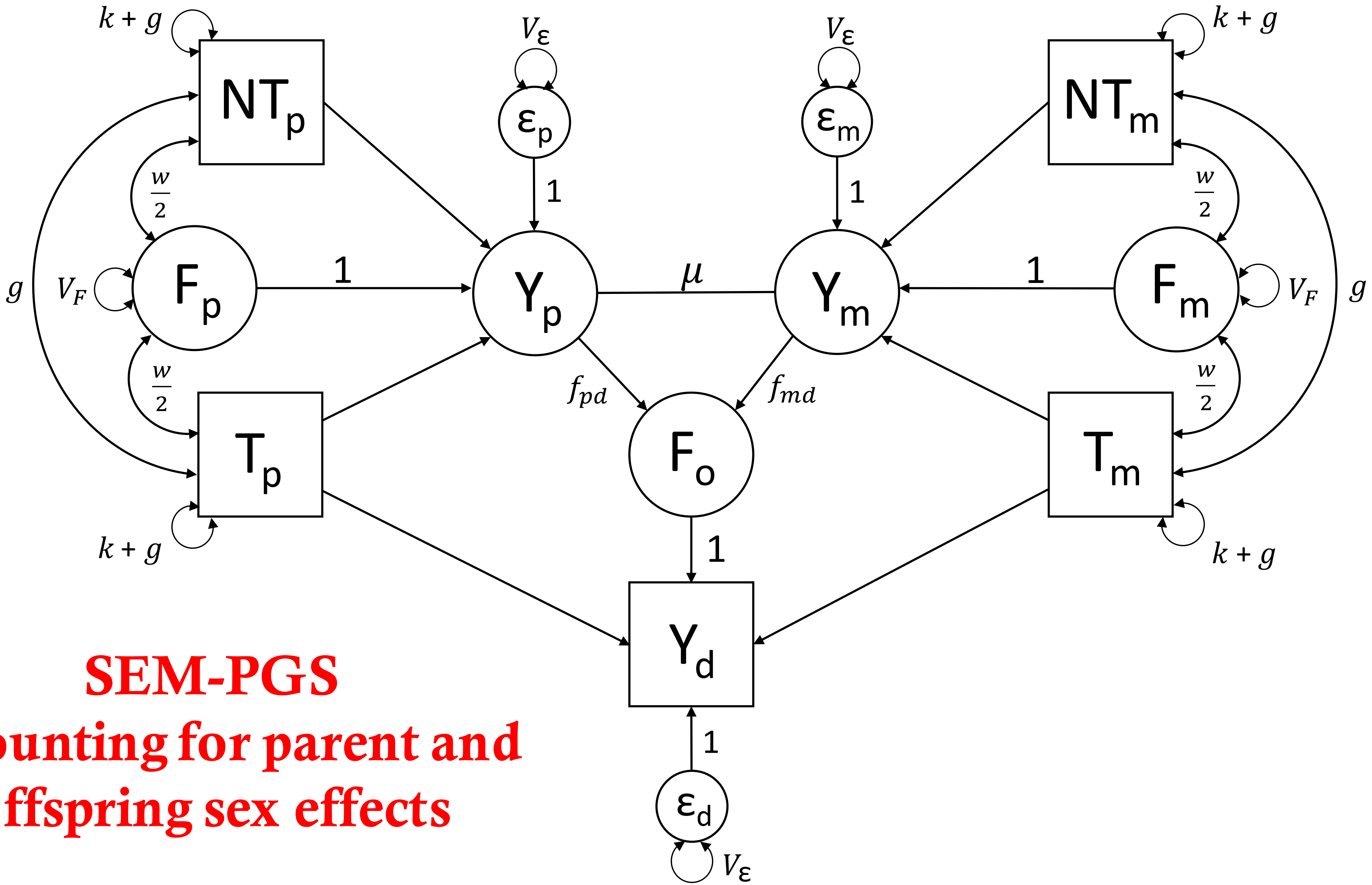
Sibling



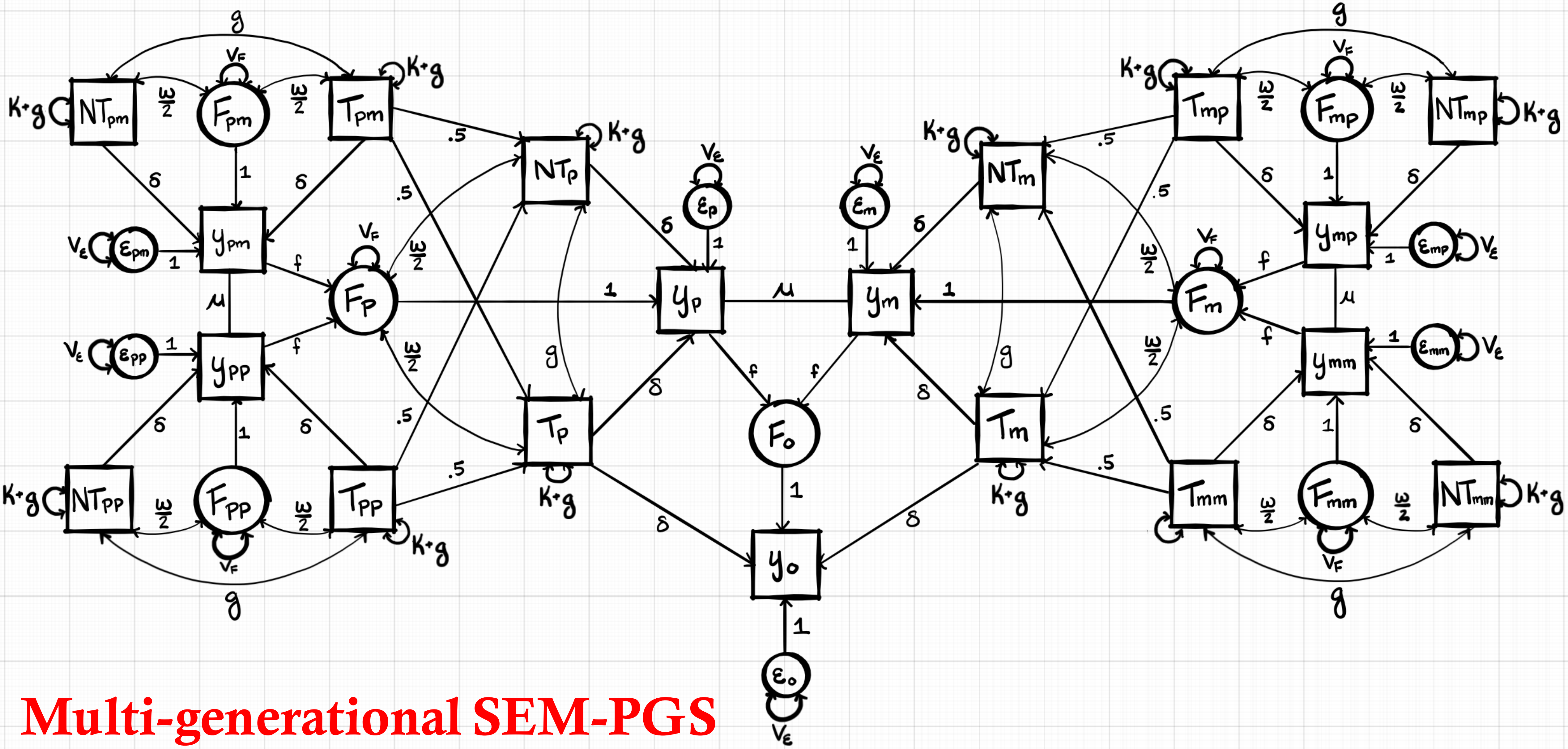
Bivariate version of SEM-PGS



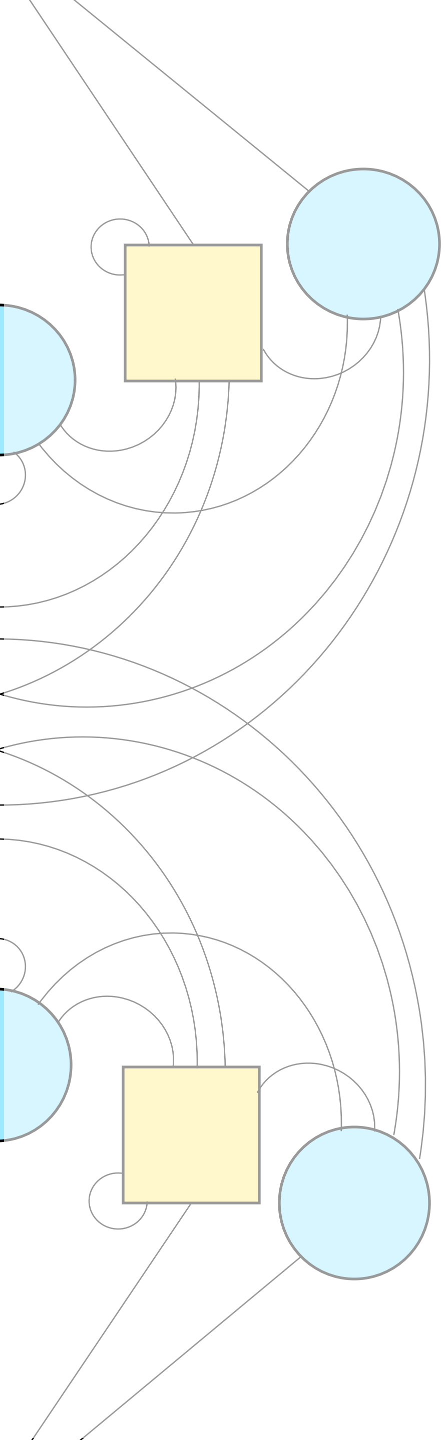
Bivariate version of SEM-PGS



SEM-PGS accounting for parent and offspring sex effects



Multi-generational SEM-PGS



Thank you!